

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 05:28:16 ; Search time 10901 Seconds

(without alignments)  
11204.538 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818

Sequence: 1 agagaccatcatcattatga.....aaaaaaaaaaaaaaaaa 2818

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seg length: 0

Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
GenBank:
1: gb_ba:*
2: gb_hvg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vt:*
30: em_hvg_hum:*
31: em_hvg_inv:*
32: em_hvg_other:*
33: em_hvg_mus:*
34: em_hvg_pin:*
35: em_hvg_rod:*
36: em_hvg_mam:*
37: em_hvg_vrt:*
38: em_sy:*
39: em_hugo_hum:*
40: em_hugo_mus:*
41: em_hugo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	2818	100.0	2818	6	BD016833
2	2791	99.0	2791	9	AB040431
3	2174.6	77.2	6564	6	BD016835
4	2174.6	77.2	11204	6	BD016860
5	2174.6	77.2	11204	9	AB040430
6	2174.6	77.2	71132	9	AC092184
7	2172	77.1	21132	6	BD016840
8	1818.6	64.5	1828	9	BC006296
9	603.4	21.4	2440	6	BD016828
10	603.4	21.4	2440	10	AF132979
11	595.4	21.1	597	12	AF529828
12	595	21.1	597	9	BT007402
13	595	21.1	597	12	BT008226
14	594.4	21.1	596	9	AF529815
15	594.4	21.1	596	9	AF529816
16	594.4	21.1	596	9	AF529815
17	594.4	21.1	596	9	AF529820
18	594.4	21.1	596	9	AF529821
19	594.4	21.1	596	9	AF529822
20	594.4	21.1	596	9	AF529823
21	594.4	21.1	596	9	AF529824
22	594.4	21.1	596	9	AF529825
23	594.4	21.1	596	9	AF529826
24	594.4	21.1	596	9	AF529827
25	594.4	21.1	596	12	AF529829
26	594.4	21.1	596	12	AF529830
27	594.4	21.1	596	12	AF529831
28	594.4	21.1	596	12	AF529832
29	594.4	21.1	596	12	AF529833
30	594.4	21.1	596	12	AF529837
31	594.4	21.1	596	12	AF529839
32	593.4	21.1	595	12	AF529834
33	592.8	21.0	596	9	AF529817
34	592.8	21.0	596	9	AF529818
35	592.8	21.0	596	12	AF529832
36	592.8	21.0	596	12	AF529836
37	592.8	21.0	596	12	AF529838
38	592.2	21.0	597	12	AF529841
39	592.2	21.0	597	12	AF529842
40	592.2	21.0	597	12	AF529843
41	592.2	21.0	597	12	AF529844
42	592.2	21.0	597	12	AF529845
43	592.2	21.0	597	12	AF529846
44	592.2	21.0	597	12	AF529847
45	592.2	21.0	597	12	AF529848

## ALIGNMENTS

```
RESULT 1
LOCUS      BD016833
DEFINITION Novel cytidine deaminase.
ACCESSION  BD016833.1 GI:22558009
VERSION    BD016833.1
KEYWORDS   JP 2001245669-A/6.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 2818)
AUTHORS   Honjo, T. and Muramatsu, M.
TITLE     Novel cytidine deaminase
JOURNAL   Patent: JP 2001245669-A 6 11-SEP-2001;
```

COMMENT  
JAPAN TOBACCO INC./TASUKU HONDO  
OS Homo sapiens (human)  
PN JP 2001245669-A/6  
PD 11-SEP-2001  
PF 28-MAR-2000 JP 2000092981  
PI TASUKU HONDO, MASAMICHI MURAMATSU  
PC C12N15/09, A61K39/395, A61K39/395, A61P11/00, A61P11/06, A61P13/12,  
PC A61P17/00,  
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC  
C12N1/21  
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08// (C12N1/21, C12R1.19), PC  
(C12N5/10, C12R1.91), C12N15/00, C12N5/00, (C12N5/00, C12R1.91) CC  
FH Key Location/Qualifiers  
FT 5'UTR (1) . . (79)  
FT CDS (80) . . (676)  
FT 3'UTR (677) . . (2818).  
FEATURES  
Source Location/Qualifiers  
1. . 2818  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 100.0%; Score 2818; DB 6; Length 2818;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGGAGACCATCATTAATTGAAGTGAATTTTCTGCTGAGACCTTGACGGAGGCGAAGA 60  
Db 1 AGGAGACCATCATTAATTGAAGTGAATTTTCTGCTGAGACCTTGACGGAGGCGAAGA 60  
QY 61 AGACACCTGGAGACCACTATGAGACAGCCCTTGATGAACCGAGGAAGTTCTTTACCA 120  
Db 61 AGACACCTGGAGACCACTATGAGACAGCCCTTGATGAACCGAGGAAGTTCTTTACCA 120  
QY 121 ATTCAAAATGTCCGCTGGGCTTAAGGATCGGCGTGAACCTTACTGTACTAGTATGA 180  
Db 121 ATTCAAAATGTCCGCTGGGCTTAAGGATCGGCGTGAACCTTACTGTACTAGTATGA 180  
QY 121 ATTCAAAATGTCCGCTGGGCTTAAGGATCGGCGTGAACCTTACTGTACTAGTATGA 180  
Db 121 ATTCAAAATGTCCGCTGGGCTTAAGGATCGGCGTGAACCTTACTGTACTAGTATGA 180  
QY 181 GAGGCGTGACAGTGTACATCCTTTCACTGGACTTTGGTTATCTTCGATATGAACG 240  
Db 181 GAGGCGTGACAGTGTACATCCTTTCACTGGACTTTGGTTATCTTCGATATGAACG 240  
QY 181 GAGGCGTGACAGTGTACATCCTTTCACTGGACTTTGGTTATCTTCGATATGAACG 240  
Db 181 GAGGCGTGACAGTGTACATCCTTTCACTGGACTTTGGTTATCTTCGATATGAACG 240  
QY 241 CTGCGACAGTGAATGTCTCTTCCGCTACATCTTGGAACCTTGGAACCTTGGAAC 300  
Db 241 CTGCGACAGTGAATGTCTCTTCCGCTACATCTTGGAACCTTGGAACCTTGGAAC 300  
QY 301 CTGCTACCGGCTACCGTGGTTACCTCTGGAAGCCCTGCTACGACTGTGCCGACATGT 360  
Db 301 CTGCTACCGGCTACCGTGGTTACCTCTGGAAGCCCTGCTACGACTGTGCCGACATGT 360  
QY 301 CTGCTACCGGCTACCGTGGTTACCTCTGGAAGCCCTGCTACGACTGTGCCGACATGT 360  
Db 301 CTGCTACCGGCTACCGTGGTTACCTCTGGAAGCCCTGCTACGACTGTGCCGACATGT 360  
QY 361 GGGCGACTTTCTGCGAGGGAACCCCAACCTCAGTCTGAGGATCTTACCGCGGCTCTA 420  
Db 361 GGGCGACTTTCTGCGAGGGAACCCCAACCTCAGTCTGAGGATCTTACCGCGGCTCTA 420  
QY 421 CTCTGTGAGAGCCGCAAGGCTGAGCCGAGGGGCTGCGCGGCTGCAACCGCGGGGT 480  
Db 421 CTCTGTGAGAGCCGCAAGGCTGAGCCGAGGGGCTGCGCGGCTGCAACCGCGGGGT 480  
QY 481 GCAATATAGCATCATGACCTTCAAGATTAATTTTACTGCTGGAATACTTTTGTGAAA 540  
Db 481 GCAATATAGCATCATGACCTTCAAGATTAATTTTACTGCTGGAATACTTTTGTGAAA 540  
QY 541 CCATGAAAGAACTTTCAAAAGCTGGGAAGGCTGCATGAAATTCAGTTCCTCCAG 600  
Db 541 CCATGAAAGAACTTTCAAAAGCTGGGAAGGCTGCATGAAATTCAGTTCCTCCAG 600  
QY 601 ACAAGCTTGGCGGATCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGAGCATTTG 660  
Db 601 ACAAGCTTGGCGGATCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGAGCATTTG 660  
QY 661 TACTTTGGGACTTTGATAGCACTTCCAGGAATGTACACAGATGAATATCTCTGCTG 720  
Db 661 TACTTTGGGACTTTGATAGCACTTCCAGGAATGTACACAGATGAATATCTCTGCTG 720

Db 661 TACTTTGGGACTTTGATAGCACTTCCAGGAATGTACACAGATGAATATCTCTGCTG 720  
QY 721 AAGACAGTGAATAAAAAAGCTCTTCAATCTTCTCTGTTTATTTCTTCAACTCTCAC 780  
Db 721 AAGACAGTGAATAAAAAAGCTCTTCAATCTTCTCTGTTTATTTCTTCAACTCTCAC 780  
QY 781 TTTCTTGAAGTTTACAGAAAAATTTATATAGACTCTTTAAAAAGATCTATGCTTGG 840  
Db 781 TTTCTTGAAGTTTACAGAAAAATTTATATAGACTCTTTAAAAAGATCTATGCTTGG 840  
QY 841 AAAATAGAGAGAGAACACAGCTCTGCGACAGGACGCTGCAATTTGGTGAAGTTTGAAT 900  
Db 841 AAAATAGAGAGAGAACACAGCTCTGCGACAGGACGCTGCAATTTGGTGAAGTTTGAAT 900  
QY 901 GCAACATTTGCTCCCTATCTGGGAATTAACAGAACTGACGAGACCTTGAGCATCTTAAAGTGT 960  
Db 901 GCAACATTTGCTCCCTATCTGGGAATTAACAGAACTGACGAGACCTTGAGCATCTTAAAGTGT 960  
QY 961 CAACGTTTTCTATAGACTTTTAGTGAAGTGAAGAGAGAGATGATCCATAAAGCATG 1020  
Db 961 CAACGTTTTCTATAGACTTTTAGTGAAGTGAAGAGAGAGATGATCCATAAAGCATG 1020  
QY 1021 GTGAGAGATCAAAATGTTTATATCAACATCCTTATTTATTTGATTCATTTGATTAAC 1080  
Db 1021 GTGAGAGATCAAAATGTTTATATCAACATCCTTATTTATTTGATTCATTTGATTAAC 1080  
QY 1081 AGTGTGTATGATATAGATTTTCTATTTCTTTCCCTTGAAGTTTACATTTCAAGTAAC 1140  
Db 1081 AGTGTGTATGATATAGATTTTCTATTTCTTTCCCTTGAAGTTTACATTTCAAGTAAC 1140  
QY 1141 AAATCTTTCATCAGGACATGATCTATAGACCTCTTAATGAGATATCGGATGTTGT 1200  
Db 1141 AAATCTTTCATCAGGACATGATCTATAGACCTCTTAATGAGATATCGGATGTTGT 1200  
QY 1201 GACCCCAACATCTCTCCAAAGCATTAATATCAATCATGCGCTGATGTTTAAATCAG 1260  
Db 1201 GACCCCAACATCTCTCCAAAGCATTAATATCAATCATGCGCTGATGTTTAAATCAG 1260  
QY 1261 CAGAGCATGTTTATATGTTTATGTAACAAGAAATGTTATGGGTGGATGAGAGTATA 1320  
Db 1261 CAGAGCATGTTTATATGTTTATGTAACAAGAAATGTTATGGGTGGATGAGAGTATA 1320  
QY 1321 GACCATCATGCTACCTTCAAGCTTAAATTAAGATCTTAAATATGCGCGAGAGAC 1380  
Db 1321 GACCATCATGCTACCTTCAAGCTTAAATTAAGATCTTAAATATGCGCGAGAGAC 1380  
QY 1381 TGTGAACAACAACCCCTAATTAATGGGTGATGTCGTAAGTGAATCTTCTGGAACCG 1440  
Db 1381 TGTGAACAACAACCCCTAATTAATGGGTGATGTCGTAAGTGAATCTTCTGGAACCG 1440  
QY 1441 AAATCTTTTAAAGAAAGTCCCTAATTTAGAAACCAACCAACTTCAATATCATATTA 1500  
Db 1441 AAATCTTTTAAAGAAAGTCCCTAATTTAGAAACCAACCAACTTCAATATCATATTA 1500  
QY 1501 GCAAAATTTGAAGAAAGTCTTGAATGTTGGGAGAGAAATCTATTTGGCTCTCGT 1560  
Db 1501 GCAAAATTTGAAGAAAGTCTTGAATGTTGGGAGAGAAATCTATTTGGCTCTCGT 1560  
QY 1561 GGGTCTTTATCTCAGAAATGCGCAATCAAGTCTAAGATTTGCTATCTTTGTATGTGTGT 1620  
Db 1561 GGGTCTTTATCTCAGAAATGCGCAATCAAGTCTAAGATTTGCTATCTTTGTATGTGTGT 1620  
QY 1621 GATGCTTCTCCAAAGGATATTAATCATATAGAGTGTGACAAACAGAAATGATTA 1680  
Db 1621 GATGCTTCTCCAAAGGATATTAATCATATAGAGTGTGACAAACAGAAATGATTA 1680  
QY 1681 AGTGTGAACCGTGGACACAGCTCATATGTTCTTGAAGGTTGAGAGGAGAGGA 1740  
Db 1681 AGTGTGAACCGTGGACACAGCTCATATGTTCTTGAAGGTTGAGAGGAGAGGA 1740  
QY 1741 TGGCTTGAACAACAGTGTTCAGAGGCAAGCTGGGCAACATTAACAGATCTGTCTCAA 1800  
Db 1741 TGGCTTGAACAACAGTGTTCAGAGGCAAGCTGGGCAACATTAACAGATCTGTCTCAA 1800

```

QY 1801 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGCGTGTGGCTCAAGCCTGTATATCC 1860
Db 1801 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGCGTGTGGCTCAAGCCTGTATATCC 1860
QY 1861 CAGACCTTTGGAGGCGCGCGCGGATCACCCTGTGTGTGAGAGTTGAGACCGCT 1920
Db 1861 CAGACCTTTGGAGGCGCGCGCGGATCACCCTGTGTGTGAGAGTTGAGACCGCT 1920
QY 1921 GCGCAATGAGCAAAACCCGCTGTATCAAAATGCAAAATTAAGCAGGCGGTGTAC 1980
Db 1921 GCGCAATGAGCAAAACCCGCTGTATCAAAATGCAAAATTAAGCAGGCGGTGTAC 1980
QY 1981 AGGACCTGTATATCCAGCTACTTGGAGGCTGAGGAGAGAGATGCTTGAACCA 2040
Db 1981 AGGACCTGTATATCCAGCTACTTGGAGGCTGAGGAGAGAGATGCTTGAACCA 2040
QY 2041 GGTGAGGTTGCAAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 GGTGAGGTTGCAAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 AGACTGTCTCGAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db 2101 AGACTGTCTCGAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 2161 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Db 2161 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2221 GCCAATGAGGAGTCCCTATTTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGT 2280
Db 2221 GCCAATGAGGAGTCCCTATTTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGT 2280
QY 2281 GACAGTGAAGAAATATTCAGAAATACCATATCCCTGCGCTGATTAACCTGCA 2340
Db 2281 GACAGTGAAGAAATATTCAGAAATACCATATCCCTGCGCTGATTAACCTGCA 2340
QY 2341 TGCATGAGAGATGAGAGAGATCCACAGAGAAACCTGATGACACATGCTTAT 2400
Db 2341 TGCATGAGAGATGAGAGAGATCCACAGAGAAACCTGATGACACATGCTTAT 2400
QY 2401 TTATGTACATAGTTTGTAAAAGAGTAAAAATGTATCTTCAATGATTAAT 2460
Db 2401 TTATGTACATAGTTTGTAAAAGAGTAAAAATGTATCTTCAATGATTAAT 2460
QY 2461 TTATATATTTTGGCTGCTAATGATTTTATTAACATGATTTCTTTCTGAT 2520
Db 2461 TTATATATTTTGGCTGCTAATGATTTTATTAACATGATTTCTTTCTGAT 2520
QY 2521 AATGAGTCTCAAGTCTCAATAATTTATTAACCTTGAATGATTTCTTAAC 2580
Db 2521 AATGAGTCTCAAGTCTCAATAATTTATTAACCTTGAATGATTTCTTAAC 2580
QY 2581 GTATATGTAACATGAGTGAATGCTGCTGAGAGCAATTTCTGATTTTGAAT 2640
Db 2581 GTATATGTAACATGAGTGAATGCTGCTGAGAGCAATTTCTGATTTTGAAT 2640
QY 2641 TTATATGAGCAAAATTTGCTTGGCTCACTTCAATCAATTAATTAATTAAT 2700
Db 2641 TTATATGAGCAAAATTTGCTTGGCTCACTTCAATCAATTAATTAATTAAT 2700
QY 2701 TTTGAGAGTGTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760
Db 2701 TTTGAGAGTGTGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760
QY 2761 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2818
Db 2761 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2818

```

RESULT 2  
AB040431  
LOCUS AB040431 2791 bp mRNA linear PRI 03-OCT-2000

```

DEFINITION Homo sapiens AID mRNA for activation-induced cytidine deaminase,
complete CDS.
ACCESSION AB040431
VERSION AB040431.1 GI:9988409
KEYWORDS AID; activation-induced cytidine deaminase; Human AID.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (sites)
REFERENCE Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo,T.
AUTHORS Isolation, tissue distribution, and chromosomal localization of the
TITLE human activation-induced cytidine deaminase (Aid) gene
JOURNAL Genomics 68 (1), 85-88 (2000)
MEDLINE 20408890
PUBMED 10950930
REFERENCE
AUTHORS
2 (sites)
Revy,P., Muto,T., Levy,Y., Geisemann,F., Plebani,A., Sanal,O.,
Catalan,N., Forveille,M., Dufourcq-Lagelouse,R., Gennery,A.,
Tazcan,I., Eszy,F., Kayserili,H., Ugazio,A.G., Brousse,N.,
Muramatsu,M., Notarangelo,L.D., Kinoshita,K., Honjo,T., Fischer,A.
and Durandy,A.
Activation-induced cytidine deaminase (Aid) deficiency causes the
autosomal recessive form of the Hyper-IgM syndrome (HIGM2)
Cell 102 (5), 565-575 (2000)
JOURNAL 20460541
MEDLINE 11007475
PUBMED
REFERENCE
AUTHORS
3 (bases 1 to 2791)
Muto,T., Muramatsu,M., Taniwaki,M., Kinoshita,K. and Honjo,T.
Direct Submission
JOURNAL Submitted (18-MAR-2000) Tasuku Honjo, Kyoto University, Department
of Medical Chemistry, Faculty of Medicine, Yoshida, Sakyo-ku,
Kyoto, Kyoto 606-8501, Japan (E-mail:honjo@med.kyoto-u.ac.jp,
Tel:81-75-753-4371 (ex.4371), Fax:81-75-753-4388)
FEATURES
location/Qualifiers
1..2791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2791
/gene="AID"
77..673
/gene="AID"
/codon_start=1
/product="activation-induced cytidine deaminase"
/protein_id="BAI2721.1"
/db_xref="GI:9988410"
/translation="MPSLMNRKRPKYGRNRMAGRETYCYVGRDSTFSL
DFSLRNKGCHVELFLKRTISDMDLPKRCYRVFTSMSPCYDCAHVADLRNP
NLSLRFTALVFCEDRKAEPEGLRLHAGVQIALMTPEKDYCWNTVEVNERFEK
AWGLHENSVRSLRQLRILLPLYEVDDLRDAFRTGL"
ORIGIN
Query Match 99.0%; Score 2791; DB 9; Length 2791;
Bee Local Similarity 100.0%; Pred.No. 0;
Matches 2791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 GAACCATCAATTAATGAGAGATTTTCTGCGCTGAGCTGAGGAGGAGGAGGAGGAG 63
Db 1 GAACCATCAATTAATGAGAGATTTTCTGCGCTGAGCTGAGGAGGAGGAGGAGGAG 60
QY 64 CACTGTGACACCACTATGAGCAGCCTCTGTATGAGACCGAGAGAGTTTCTTACCAATT 123
Db 61 CACTGTGACACCACTATGAGCAGCCTCTGTATGAGACCGAGAGAGTTTCTTACCAATT 120
QY 124 CAATAATGTCCGTGGCTTAAGGCTGCGGTGAGACCTACTGTCTACGTATGAGAG 183
Db 121 CAATAATGTCCGTGGCTTAAGGCTGCGGTGAGACCTACTGTCTACGTATGAGAG 180
QY 184 GCGTGACAGTGCATACCTCTTCACTGACCTTGTGTATCTTGGCAATTAAGACGGCTG 243
Db 181 GCGTGACAGTGCATACCTCTTCACTGACCTTGTGTATCTTGGCAATTAAGACGGCTG 240

```

244 CCACGCGAATGCTCTTCCCTCCGCTACATCTCGGACTGGGACCTAGACCTTGCGCTG 303  
241 CCAAGTGGAAATGCTCTTCCCTCCGCTACATCTCGGACTGGGACCTAGACCTTGCGCTG 300  
304 CTACCGGCTCACTCGGTTCACTCTCTGGAGCCCTGCTACAGCTGTCGCCAGATGGG 363  
301 CTACCGGCTCACTCGGTTCACTCTCTGGAGCCCTGCTACAGCTGTCGCCAGATGGG 360  
364 CGACTTCTGCGAGGGAACCCCAACCTCAGTCTGAGGATCTGACCGCGCGCTCACT 423  
361 CGACTTCTGCGAGGGAACCCCAACCTCAGTCTGAGGATCTGACCGCGCGCTCACT 420  
424 CTGTAGAGACCGCAAGGCTGAGCCCGAGGGGCTGCGCGGCTGACACCGCGCGGATGA 483  
421 CTGTAGAGACCGCAAGGCTGAGCCCGAGGGGCTGCGCGGCTGACACCGCGCGGATGA 480  
484 AATAGCCATCTAGACCTTCAAAAGATTTATTTTACTGCTGGAATATCTTTGTAGAAAACA 543  
481 AATAGCCATCTAGACCTTCAAAAGATTTATTTTACTGCTGGAATATCTTTGTAGAAAACA 540  
544 TGAAGAACTTTCAAGCTGGGGAAGGGCTGCAATGGAATAATTCAGTCTGCTCCAGACA 603  
541 TGAAGAACTTTCAAGCTGGGGAAGGGCTGCAATGGAATAATTCAGTCTGCTCCAGACA 600  
604 GCTTCGGGCACTCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGCAATTTGCTAC 663  
601 GCTTCGGGCACTCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGCAATTTGCTAC 660  
664 TTTGGGACTTTGATAGCACTTCAGGAATGTCAACACAGATGAATATCTCTGTGAAG 723  
661 TTTGGGACTTTGATAGCACTTCAGGAATGTCAACACAGATGAATATCTCTGTGAAG 720  
724 ACAGTGAATAAAAAACAGTCTCTCAAGTCTTCTGTTTATTTCTTCAACTCTCACTTT 783  
721 ACAGTGAATAAAAAACAGTCTCTCAAGTCTTCTGTTTATTTCTTCAACTCTCACTTT 780  
784 CTAGAGTTTACAGAAAAATTTATATAGACTCTTTAAAGATCTATGCTGAGAA 843  
781 CTAGAGTTTACAGAAAAATTTATATAGACTCTTTAAAGATCTATGCTGAGAA 840  
844 ATAGAGAGAGAACACAGGCTGCGCCAGGAGAGTGTGCAATTTGCTCAAGTTTGAATGA 903  
841 ATAGAGAGAGAACACAGGCTGCGCCAGGAGAGTGTGCAATTTGCTCAAGTTTGAATGA 900  
904 ACATTGTCCTCTACTGGAATTAACAGACTGCAAGGACTTGGAGACCTCTAAAGTGTCA 963  
901 ACATTGTCCTCTACTGGAATTAACAGACTGCAAGGACTTGGAGACCTCTAAAGTGTCA 960  
964 CGTTTTCATAGACTTTTAGTAGAGATGAGAGCAAGAGTGAATCTCTAAAGATGATG 1023  
961 CGTTTTCATAGACTTTTAGTAGAGATGAGAGCAAGAGTGAATCTCTAAAGATGATG 1020  
1024 AGAGAGATCAATGTTTATATCAACATCTTTATTTATTTGATTTCAATTTGATTAACAG 1083  
1021 AGAGAGATCAATGTTTATATCAACATCTTTATTTATTTGATTTCAATTTGATTAACAG 1080  
1084 GGTGTAGTAGATGATTTTCTATTTCTTTCCCTTGAAGTTTCACTTTCAAGTAACACAA 1143  
1081 GGTGTAGTAGATGATTTTCTATTTCTTTCCCTTGAAGTTTCACTTTCAAGTAACACAA 1140  
1144 CTCTTCATCAGGCGATGATCTATAGACTCTCTAATGAGAGTATCTGGGTGATTTGAC 1203  
1141 CTCTTCATCAGGCGATGATCTATAGACTCTCTAATGAGAGTATCTGGGTGATTTGAC 1200  
1204 CCCAAGCATCTCTCCAAAGCATTAATATCAATCATGCGCTGTATGTTTATCTACGAG 1263  
1201 CCCAAGCATCTCTCCAAAGCATTAATATCAATCATGCGCTGTATGTTTATCTACGAG 1260  
1264 AAGCATGTTTATGTTTGTACAAAGAGATTTGATGGGTGGGATGAGAGTATAGAC 1323  
1261 AAGCATGTTTATGTTTGTACAAAGAGATTTGATGGGTGGGATGAGAGTATAGAC 1320  
1324 CATGATGTCACCTTCAAGCTACTTTAATTAAGATCTTTAAATGGGACGAGAGCTGT 1383

1321 CATGATGTCACCTTCAAGCTACTTTAATTAAGATCTTTAAATGGGACGAGAGCTGT 1380  
1384 GAAACAAGACCCCTAATTAATGGGTTGATGTTGAAGTACCAATCTTTGGAAACGCAAA 1443  
1381 GAAACAAGACCCCTAATTAATGGGTTGATGTTGAAGTACCAATCTTTGGAAACGCAAA 1440  
1444 CTCTTTAAGGAGCCCTAATTTAGAAACCCCAAACTTGACATATCAATTAATGCA 1503  
1441 CTCTTTAAGGAGCCCTAATTTAGAAACCCCAAACTTGACATATCAATTAATGCA 1500  
1504 AACATTTGAAGAAAGTTGCTTGAATGTTGGGAGAGAAATCTATGCTCTGCGGG 1563  
1501 AACATTTGAAGAAAGTTGCTTGAATGTTGGGAGAGAAATCTATGCTCTGCGGG 1560  
1564 TCTCTTCACTCAGAAATGCCAATCAGGTCAGGTTGCTACATTTTGTATGTGTGAT 1623  
1561 TCTCTTCACTCAGAAATGCCAATCAGGTCAGGTTGCTACATTTTGTATGTGTGAT 1620  
1624 GCTTCTCCCAAGGATATTTAATCTATTAAGAGTTGTGACAAACAGAAATGATAAGC 1683  
1621 GCTTCTCCCAAGGATATTTAATCTATTAAGAGTTGTGACAAACAGAAATGATAAGC 1680  
1684 TGGCAACCGTGGCAACAGCTCATAGTTCTAGCTCTTGGAGGTTGAGAGAGGATGG 1743  
1681 TGGCAACCGTGGCAACAGCTCATAGTTCTAGCTCTTGGAGGTTGAGAGAGGATGG 1740  
1744 CTGAAACAAGTGTTCAGAGGCGAGCTGGCAACATTAACAGATCTGTCTCAAAA 1803  
1741 CTGAAACAAGTGTTCAGAGGCGAGCTGGCAACATTAACAGATCTGTCTCAAAA 1800  
1804 AAAAAAAAAAAAAAAAAAGAAAGAGAGGCGCGGCTGTGCTACGCTGTATCTCCAG 1863  
1801 AAAAAAAAAAAAAAAAAAGAAAGAGAGGCGCGGCTGTGCTACGCTGTATCTCCAG 1860  
1864 CACTTGGAGGCGGAGCGGCGGAGTCACTGTGCTCAGAGTTGAGACCGAGCTGGC 1923  
1861 CACTTGGAGGCGGAGCGGCGGAGTCACTGTGCTCAGAGTTGAGACCGAGCTGGC 1920  
1924 CAACATGGCAAAACCCGCTGTGTACTCAAAATGCAAAATTTAGCCAGGCTGTAGCAGG 1983  
1921 CAACATGGCAAAACCCGCTGTGTACTCAAAATGCAAAATTTAGCCAGGCTGTAGCAGG 1980  
1984 CACTGTATCTCCAGCTTACTTGGAGGCTGAGGCAAGAGATGCTTGAACCCAGAGGT 2043  
1981 CACTGTATCTCCAGCTTACTTGGAGGCTGAGGCAAGAGATGCTTGAACCCAGAGGT 2040  
2044 GAGGTTGCAATAGCTGAGATCGTGCCGTTGCACTCAGCCTGGGCGGACAAAGCAAGA 2103  
2041 GAGGTTGCAATAGCTGAGATCGTGCCGTTGCACTCAGCCTGGGCGGACAAAGCAAGA 2100  
2104 CTCTGTCTCAAAAAAATTTTAAAAAAAAAGAGAGAGAGAAAGAGAAATTTGGGA 2163  
2101 CTCTGTCTCAAAAAAATTTTAAAAAAAAAGAGAGAGAGAAAGAGAAATTTGGGA 2160  
2164 GAGAGAGTGGGGAAGATTCAGAGAAATTTGCTTTATCAACAAATGTAGAGGCC 2223  
2161 GAGAGAGTGGGGAAGATTCAGAGAAATTTGCTTTATCAACAAATGTAGAGGCC 2220  
2224 AATAAGGATCTCCATTTGTCTCTTTGGTGTCTAATTTGCTTAACTGTCTTGAC 2283  
2221 AATAAGGATCTCCATTTGTCTCTTTGGTGTCTAATTTGCTTAACTGTCTTGAC 2280  
2284 AGTGAAGAAATTTATCAAGATTAACATATCCCTGTGCGCTTATTAAGCAACCTTGC 2343  
2281 AGTGAAGAAATTTATCAAGATTAACATATCCCTGTGCGCTTATTAAGCAACCTTGC 2340  
2344 AATGAAGATGACAGATCCACAGAAACTTGAATGACAACTGTCTTATTTATCTTA 2403  
2341 AATGAAGATGACAGATCCACAGAAACTTGAATGACAACTGTCTTATTTATCTTA 2400  
2404 TTGTACATAGTTTGTAAAGATTTAAATTTGTTACTTCAATGATTTATTTTAA 2463



Db 2401 TTGTACATTAAGTTGTAAGTAAAGTTAAATGTTACTCATGTATTCATTATATTGA 2460  
Qy 2464 TATTATTTGGCTGTAATGATTTTATTAACATGATTTCTTTGATATATGAAAT 2523  
Db 2461 TATTATTTGGCTGTAATGATTTTATTAACATGATTTCTTTGATATATGAAAT 2520  
Qy 2524 GGAGCTCAGAGCTCTCAATTAATTAATTAACATGATTTCTTTGATATATGAAAT 2583  
Db 2521 GGAGCTCAGAGCTCTCAATTAATTAATTAACATGATTTCTTTGATATATGAAAT 2580  
Qy 2584 ATTGTACATTAAGTTGTAAGTAAAGTTAAATGTTACTCATGTATTCATTATATTGA 2643  
Db 2581 ATTGTACATTAAGTTGTAAGTAAAGTTAAATGTTACTCATGTATTCATTATATTGA 2640  
Qy 2644 ATGACAGCAATTTGCTTGGCTCACTTCATTCAGTTAAATTAATTAATTAATTTT 2703  
Db 2641 ATGACAGCAATTTGCTTGGCTCACTTCATTCAGTTAAATTAATTAATTAATTTT 2700  
Qy 2704 GGAAGCTGTAGAGTAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTA 2763  
Db 2701 GGAAGCTGTAGAGTAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTA 2760  
Qy 2764 AATTAATAATCAATGATATGATGAAATTAATTAATTAATTAATTAATTAATTAATTA 2794  
Db 2761 AATTAATAATCAATGATATGATGAAATTAATTAATTAATTAATTAATTAATTAATTA 2791

RESULT 3  
BD016835  
LOCUS BD016835  
DEFINITION Novel cytidine deaminase.  
ACCESSION BD016835  
VERSION BD016835.1  
KEYWORDS JP 2001245669-A/8.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 6564)  
AUTHORS Honjo, T. and Muramatsu, M.  
TITLE Novel cytidine deaminase  
JOURNAL Patent: JP 2001245669-A 8.11-SEP-2001;  
JAPAN TOBACCO INC. TASUKU HONJO  
COMMENT OS Homo sapiens (human)  
PN JP 2001245669-A/8  
PD 11-SEP-2001  
PI 28-MAR-2000 JP 2000092981  
PI TASUKU HONJO, MASAMICHI MURAMATSU  
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P11/06, A61P13/12,  
PC A61P17/00,  
PC A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12N9/78, C12P21/02, C12P21/08, C12N1/21, C12R1.19), PC  
(C12N5/10, C12R1.91), C12N15/00, C12N5/00, C12N5/00, C12R1.91) CC  
FH Key Location/Qualifiers  
1. 6564  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 77.2%; Score 2174.6; DB 6; Length 6564;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 617 CTTTGGCTGTATGAGTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 676  
Db 3735 CTTCAGGCTGTATGAGTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 3794  
Qy 677 TAGCACTTCAGAGTATGACACAGATGAATATCTGCTGAGAGACAGTGAATTAATTA 736  
Db 3795 TAGCACTTCAGAGTATGACACAGATGAATATCTGCTGAGAGACAGTGAATTAATTA 3854

Qy 737 AACAGCTCGAAGCTCTCTGTTTTTATTTTCAACTCATCTTCTTGAAGTTTGA 796  
Db 3855 AACAGCTCGAAGCTCTCTGTTTTTATTTTCAACTCATCTTCTTGAAGTTTGA 3914  
Qy 797 GAAAAATATTTATTAAGCACTTTTAAAGATCTATGCTTGAAATTAAGAGAGAC 856  
Db 3915 GAAAAATATTTATTAAGCACTTTTAAAGATCTATGCTTGAAATTAAGAGAGAC 3974  
Qy 857 AACAGCTCGGCGAGGAGCTGCTGCAATGCTGAGATTTGAAATGACAACTTGTCCCTTA 916  
Db 3975 AACAGCTCGGCGAGGAGCTGCTGCAATGCTGAGATTTGAAATGACAACTTGTCCCTTA 4034  
Qy 917 CTGGGAATTAACAGAACTGACAGACTGAGAGACTCTTAAGTGTCAACGTTTTCTATGA 976  
Db 4035 CTGGGAATTAACAGAACTGACAGACTGAGAGACTCTTAAGTGTCAACGTTTTCTATGA 4094  
Qy 977 CTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1036  
Db 4095 CTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4154  
Qy 1037 TTTTATATCAACATCTTTTATTTATTTGATTCATTTGATTAAGTGTGATGATTA 1096  
Db 4155 TTTTATATCAACATCTTTTATTTATTTGATTCATTTGATTAAGTGTGATGATTA 4214  
Qy 1097 GATTTTCTATTTCTTTTCCCTTGAAGTTTACTTTCAAGTAAACAGAACTTTCCATGAG 1156  
Db 4215 GATTTTCTATTTCTTTTCCCTTGAAGTTTACTTTCAAGTAAACAGAACTTTCCATGAG 4274  
Qy 1157 CCAATGATATTAAGACCTCCCTTAATGAGATATCTGGGTGATTTGAGACCCCAACATCTC 1216  
Db 4275 CCAATGATATTAAGACCTCCCTTAATGAGATATCTGGGTGATTTGAGACCCCAACATCTC 4334  
Qy 1217 TCAGAGATTAATATCAATCATGAGCTGATTTTAAATCAAGAGATTTTAA 1276  
Db 4335 TCAGAGATTAATATCAATCATGAGCTGATTTTAAATCAAGAGATTTTAA 4394  
Qy 1277 TGTTTGTAAGAAAGATTTGATTTGAGGTGAGATGAGAGTATTAAGCATGATCTAC 1336  
Db 4395 TGTTTGTAAGAAAGATTTGATTTGAGGTGAGATGAGAGTATTAAGCATGATCTAC 4454  
Qy 1337 CTTCAAGTATTTAAATTAAGAGATCTTAATTAAGTGGCGAGAGACTGTGAACAGACACC 1396  
Db 4455 CTTCAAGTATTTAAATTAAGAGATCTTAATTAAGTGGCGAGAGACTGTGAACAGACACC 4514  
Qy 1397 TAAATAGGTTGATATGCTGAGTAAAGTAAATCTTGGAAACGAACTCTTTAAAGAA 1456  
Db 4515 TAAATAGGTTGATATGCTGAGTAAAGTAAATCTTGGAAACGAACTCTTTAAAGAA 4574  
Qy 1457 GTCCCTAATTTAGAAACACCACTTCAATATCAATATTAATTAAGCAATTTGAAGG 1516  
Db 4575 GTCCCTAATTTAGAAACACCACTTCAATATTAATTAATTAAGCAATTTGAAGG 4634  
Qy 1517 AAGTTGCTGAATGTTGGGAGAGAGAAATCTATGCTGCTGAGGCTCTTCAATCTCA 1576  
Db 4635 AAGTTGCTGAATGTTGGGAGAGAGAAATCTATGCTGCTGAGGCTCTTCAATCTCA 4694  
Qy 1577 GAATGCGCAATCAAGTGAAGTTTGTCAATTTTGTATGTTGTGATGCTTCTCCCAAG 1636  
Db 4695 GAATGCGCAATCAAGTGAAGTTTGTCAATTTTGTATGTTGTGATGCTTCTCCCAAG 4754  
Qy 1637 GTATATTAATTAATTAAGAGTGTGACAAACAGATTAAGCTGGAACCGTGGC 1696  
Db 4755 GTATATTAATTAATTAAGAGTGTGACAAACAGATTAAGCTGGAACCGTGGC 4814  
Qy 1697 ACAGCTCATATGTTTCTGCTTTGAGGTTGAGAGGAGAGAGTGTGAACAGAGT 1756  
Db 4815 ACAGCTCATATGTTTCTGCTTTGAGGTTGAGAGGAGAGAGTGTGAACAGAGT 4814  
Qy 1757 GTTCAAGGCAAGCTGAGGCAATTAACAGATCTGCTCTCAAAAAAATTAATTAATTA 1816  
Db 4875 GTTCAAGGCAAGCTGAGGCAATTAACAGATCTGCTCTCAAAAAAATTAATTAATTA 4934

LOCUS	DEFINITION	Novel	cyclidine	deaminase.	11204 bp	DNA	linear	PAT	27-AUG-2002
QY	1817	AAGAAAGAGAGAGGGCGCGGCGTGGAGGTCAAGCTGTAATCCAGACACTTTGGGAGGC	1817						
Db	4935	AAGAAAGAGAGAGGGCGCGGCGTGGAGGTCAAGCTGTAATCCAGACACTTTGGGAGGC	4935						
QY	1877	CGAGCCGGGCGGATCACTGTGTGCAGAGATTGGAGCCAGCTGGCCAAATGGCAAA	1938						
Db	4995	CGAGCCGGGCGGATCACTGTGTGCAGAGATTGGAGCCAGCTGGCCAAATGGCAAA	5054						
QY	1937	CCCCGTCTGTACTCAAAATGCAAAAATTAGCCAGGCGTGGTAGCAGGCACTGTATCC	1996						
Db	5055	CCCCGTCTGTACTCAAAATGCAAAAATTAGCCAGGCGTGGTAGCAGGCACTGTATCC	5114						
QY	1997	AGCTACTTGGGAGGCTGAGGCAAGAAATGCTTGAATCCCGAGAGTGAAGTTCAGTA	2056						
Db	5115	AGCTACTTGGGAGGCTGAGGCAAGAAATGCTTGAATCCCGAGAGTGAAGTTCAGTA	5174						
QY	2057	AGCTGAGATCGGCGCGTTGCACTCCAGCCTGGGCGACAAGAGCAAGACTGTCTCAGAA	2116						
Db	5175	AGCTGAGATCGGCGCGTTGCACTCCAGCCTGGGCGACAAGAGCAAGACTGTCTCAGAA	5234						
QY	2117	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAAACAATATTGGGAGAGAGATGGGG	2176						
Db	5235	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAAACAATATTGGGAGAGAGATGGGG	5294						
QY	2177	AAGCATTCGACAGGAATTTGTGCTTATTCACAAATGTAAAGAGCCATTAAGGATCC	2236						
Db	5295	AAGCATTCGACAGGAATTTGTGCTTATTCACAAATGTAAAGAGCCATTAAGGATCC	5354						
QY	2237	TATTTGTCTCTTTGGTGTCTATTGTGCCAACAACGTCTTGAACGTGAGAAAAATA	2296						
Db	5355	TATTTGTCTCTTTGGTGTCTATTGTGCCAACAACGTCTTGAACGTGAGAAAAATA	5414						
QY	2297	TTCAGAATTAACATATCCCTGCGCGTATTACCTAGCAACCTTGCAATGAATGAGC	2356						
Db	5415	TTCAGAATTAACATATCCCTGCGCGTATTACCTAGCAACCTTGCAATGAATGAGC	5474						
QY	2337	AGATTCACAGGAAAACTTGAATGCAACACTGCTTATTATTTAATTTGATGACTAATT	2416						
Db	5475	AGATTCACAGGAAAACTTGAATGCAACACTGCTTATTATTTAATTTGATGACTAATT	5534						
QY	2417	TGTAAAAAGTTAAAAATTGTACTCTCACTGATTCATTATTAATTTATTTTGGCT	2476						
Db	5535	TGTAAAAAGTTAAAAATTGTACTCTCACTGATTCATTATTAATTTATTTTGGCT	5594						
QY	2477	CTAATGATTTTATTAATAATGATTTCTTTCTGATATATGTAAGAGAGTCCAAAGC	2536						
Db	5595	CTAATGATTTTATTAATAATGATTTCTTTCTGATATATGTAAGAGAGTCCAAAGC	5654						
QY	2537	TTCAATAATTATAAATTAGAAAAATGATTCATATACACAGATATGTAATGTAATTCG	2596						
Db	5655	TTCAATAATTATAAATTAGAAAAATGATTCATATACACAGATATGTAATGTAATTCG	5714						
QY	2597	AGTAATGCTGTAGAGAGCCATTCTCTGTAATTTTATGTAACCTTTATGACAGCAATT	2656						
Db	5715	AGTAATGCTGTAGAGAGCCATTCTCTGTAATTTTATGTAACCTTTATGACAGCAATT	5774						
QY	2657	TGCTTCTGGCTCACTTCAATCAGTTAAATAATGATTAATAATTTGGAAGCTGTAAAG	2716						
Db	5775	TGCTTCTGGCTCACTTCAATCAGTTAAATAATGATTAATAATTTGGAAGCTGTAAAG	5834						
QY	2717	ATTAATAATCCAAATTAATAATTAATTAATTAATGATTTATGTAAGTTAAATTAATAATCAG	2776						
Db	5835	ATTAATAATCCAAATTAATAATTAATTAATTAATGATTTATGTAAGTTAAATTAATAATCAG	5894						
QY	2777	TATGATGGAATTAACCTGAAA	2797						
Db	5895	TATGATGGAATTAACCTGAGA	5915						

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal
BD016860	1	GI:22558036	JP 2001245669-A/33.	Homo sapiens (human)	Horio, T. and Muramatsu, M.	Novel cytidine deaminase	Patent: JP 2001245669-A 33	11-SEP-2001;
PC	C12N1/21	PC	A61P27/02, A61P27/16, A61P37/02, A61P37/08, C07K6/18, C12N1/19, C12N1/21, C12N5/10, C12R1:91, C12N15/00, C12N5/00, C12N5/00, C12R1:91	CC				
FEATURES	source	location/Qualifiers	key	location/Qualifiers	key	location/Qualifiers	key	location/Qualifiers
ORIGIN	1. 11204	location/Qualifiers	key	location/Qualifiers	key	location/Qualifiers	key	location/Qualifiers
Query Match	77.2%	Score 2174.6;	DB 6;	Length 11204;				
Best Local Similarity	99.8%	Pred. No. 0;						
Matches 2177;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
61.7	CTTTGCCCGCTGTAGAGTTGAGTCTTACGACCTTACGACGATTTTCGACTCTTGGGACTTGA	676						
89.50	CTACAGCCCCGTAATGAGTTGAGTCTTACGACGATTTTCGACTCTTGGGACTTGA	900.5						
67.7	TAGCACTTCGAGAAATGTCACACAGATGAAATATCTGCTGAGACAGTGGATATAA	736						
90.10	TAGCACTTCGAGAAATGTCACACAGATGAAATATCTGCTGAGACAGTGGATATAA	906.5						
73.7	AACAGCTTCGAGTCTTCTGTTTATCTTCAACTCTGACTTTCTTAGAGTTTACA	796						
90.70	AACAGCTTCGAGTCTTCTGTTTATCTTCAACTCTGACTTTCTTAGAGTTTACA	912.9						
79.7	GAAATATTTTATATATGCACTCTTTAAAGATCTATGCTTTAAATATAGAGAGAC	856						
91.30	GAAATATTTTATATATGCACTCTTTAAAGATCTATGCTTTAAATATAGAGAGAC	918.9						
85.7	ACAAGTCTGGCAGAGGAGCTGCTCAATTTGGTGAGTTTGAATGCAACATGTCCTTA	916						
91.90	ACAAGTCTGGCAGAGGAGCTGCTCAATTTGGTGAGTTTGAATGCAACATGTCCTTA	924.9						
91.7	CTGGAAATTAACAGACTGACAGGACTTGGAGCACTCTAAAGTGTCAACGTTTTCTATGA	976						
92.50	CTGGAAATTAACAGACTGACAGGACTTGGAGCACTCTAAAGTGTCAACGTTTTCTATGA	930.9						
97.7	CTTTAGTGTGAGTGAAGAGAGAGAGTATCTTAAAGATGATGAGAGATCAATG	1036.6						
93.10	CTTTAGTGTGAGTGAAGAGAGAGAGTATCTTAAAGATGATGAGAGATCAATG	936.9						
103.7	TTTTATATCAACACTCTTATATATTTGATTCATTTAGTTTACAGTGTGATGATA	1096.6						
93.70	TTTTATATCAACACTCTTATATTTGATTCATTTAGTTTACAGTGTGATGATA	942.9						
109.7	GATTTTCTATTTCTTTTCCCTTGAAGTTTCACTTCAAGTATACAACTCTTCATCAGG	1156.6						
94.30	GATTTTCTATTTCTTTTCCCTTGAAGTTTCACTTCAAGTATACAACTCTTCATCAGG	948.9						
115.7	CCATGATCTATAGGAGCTCTCTATGAGAGTATCTGGGTGATGTGAGACCCCAACATCTC	1216.6						



Kyoto, Kyoto 606-8501, Japan (E-mail: honjo@four.med.kyoto-u.ac.jp,  
Tel: 81-75-753-4371 (ex. 4371), Fax: 81-75-753-4388)  
Location/Qualifiers

FEATURES  
Source

1. 11204

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

join(521..528,6280..6427,7807..8077,8371..8486,8956..9009)

/gene="AID"

join(521..528,6280..6427,7807..8077,8371..8486,8956..9009)

/gene="AID"

/codon\_start=1

/product="activation-induced cytidine deaminase"

/protein\_id="BAB12720.1"

/db\_xref="GI:9988408"

/translation="MDSILMRKRFYQKRWKMAKRETYLTVYKRRDSATPSL  
DPIGRKNGCHVELLFARIISDWIDPFCYKRVVTFMSPCYCARVADFLKGNP  
NLSLRIFTALYFCEDRKAPESGLRIHAGVQIALMTKDYFCWNTVEVHERTFK  
AMEGLHENSVELSRQLRIILPLVEVDLDRDAFRILGL"

## ORIGIN

Query Match 77.28; Score 2174.6; DB 9; Length 11204;  
Best Local Similarity 99.88; Pred. NO. 0;  
Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 617 CTTTGGCCCGTATGAGTGTGATGACTTACGAGCACTTGGGACTTTGGAAGCTTTGA 676  
DB 8950 CTACAGCCCTGATGAGTGTGATGACTTACGAGCACTTGGGACTTTGGAAGCTTTGA 9009  
QY 677 TAGCACTTCGAGGATGTGACACAGATGAATATCTGCTGAGAGACAGTGTATATA 736  
DB 9010 TAGCACTTCGAGGATGTGACACAGATGAATATCTGCTGAGAGACAGTGTATATA 9069  
QY 737 AACAGTCTTCAAGTCTCTCTGTTTATTTTCACTCTCACTTCTTGAAGTTTACA 796  
DB 9070 AACAGTCTTCAAGTCTCTCTGTTTATTTTCACTCTCACTTCTTGAAGTTTACA 9129  
QY 797 GAAAAATATTTATATACGACTCTTTAAAAAGATCTATGCTTGAATAAGAGAGAAC 856  
DB 9130 GAAAAATATTTATATACGACTCTTTAAAAAGATCTATGCTTGAATAAGAGAGAAC 9189  
QY 857 ACAAGTCTGCGCAGGAGCGTGTCAATTTGAGTGTGATGATGATGATGATGATGATG 916  
DB 9190 ACAAGTCTGCGCAGGAGCGTGTCAATTTGAGTGTGATGATGATGATGATGATGATG 9249  
QY 917 CTGGGAATTAACAGAACTGACAGACCTGGAGCATCTTAAGTGTCAACCTTTTCTATGA 976  
DB 9250 CTGGGAATTAACAGAACTGACAGACCTGGAGCATCTTAAGTGTCAACCTTTTCTATGA 9309  
QY 977 CTTTATAGTATGATGAGAGAGAGATGATCTTAAAAAGCATGTGTGAGAGATCAATG 1036  
DB 9310 CTTTATAGTATGATGAGAGAGAGATGATCTTAAAAAGCATGTGTGAGAGATCAATG 9369  
QY 1037 TTTTATATCAATCATCTTATATTTGATCACTTTGATTAACAGTGTGTATGATGATA 1096  
DB 9370 TTTTATATCAATCATCTTATATTTGATCACTTTGATTAACAGTGTGTATGATGATA 9429  
QY 1097 GATTTTCTATCTTTTCCCTTACGCTTACTTCAAGTATCAACAACTCTTCCATCAG 1156  
DB 9430 GATTTTCTATCTTTTCCCTTACGCTTACTTCAAGTATCAACAACTCTTCCATCAG 9489  
QY 1157 CCAATGATCTATAGAGCTCTCTAATGAGATCTGGGTGATTTGAGACCCCAACATCTC 1216  
DB 9490 CCAATGATCTATAGAGCTCTCTAATGAGATCTGGGTGATTTGAGACCCCAACATCTC 9549  
QY 1217 TCCAAAGCAATTAATCAATCAATGAGCTGTATGTTTATACAGAGAGAGATGTTTATA 1276  
DB 9550 TCCAAAGCAATTAATCAATCAATGAGCTGTATGTTTATACAGAGAGAGATGTTTATA 9609  
QY 1277 TGTGTTACAAAAGAGATGTTATAGGTGGAGATGAGATGATGATGATGATGATGATG 1336  
DB 9610 TGTGTTACAAAAGAGATGTTATAGGTGGAGATGAGATGATGATGATGATGATGATG 9669

QY 1337 CTTCAAGTACTTATATTAAGATCTTAAATGGGACAGAGACCTGTGACAAAGACACC 1396  
DB 9670 CTTCAAGTACTTATATTAAGATCTTAAATGGGACAGAGACCTGTGACAAAGACACC 9729  
QY 1397 TAATATAGGTTGATGTCTGAGTACGAACTTCTGGAACGCAACTCTTTTAAGAA 1456  
DB 9730 TAATATAGGTTGATGTCTGAGTACGAACTTCTGGAACGCAACTCTTTTAAGAA 9789  
QY 1457 GTCCCTATTTGAAACACCCCAAACTTCAATATATATTAAGCAAAATTTGAAGG 1516  
DB 9790 GTCCCTATTTGAAACACCCCAAACTTCAATATATATTAAGCAAAATTTGAAGG 9849  
QY 1517 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1576  
DB 9850 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9909  
QY 1577 GAAATGCCAATCAGATCAAGGTTGCTACATTTTGTGTGTGTGTGTGTGTGTGTGTGT 1636  
DB 9910 GAAATGCCAATCAGATCAAGGTTGCTACATTTTGTGTGTGTGTGTGTGTGTGTGTGT 9969  
QY 1637 GTATATTAATTAATTAAGAGATTTGACAAAACAGATGATTAAGCTGCGAAC 1696  
DB 9970 GTATATTAATTAATTAAGAGATTTGACAAAACAGATGATTAAGCTGCGAAC 10029  
QY 1697 ACAAGCTCATATGTTCTGAGTCTGTTGGAGGTTGAGAGAGAGATGCTTGAACAAGT 1756  
DB 10030 ACAAGCTCATATGTTCTGAGTCTGTTGGAGGTTGAGAGAGAGATGCTTGAACAAGT 10089  
QY 1757 GTTCAAGGCGCCTGGGCAACATTAACAATCTGTCTCAAAAAA 1816  
DB 10090 GTTCAAGGCGCCTGGGCAACATTAACAATCTGTCTCAAAAAA 10149  
QY 1817 AAGAAAGAGAGAGGCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1876  
DB 10150 AAGAAAGAGAGAGGCGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10209  
QY 1877 CGAGCCGGGCGGATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1936  
DB 10210 CGAGCCGGGCGGATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 10269  
QY 1937 CCGCGTGTATCAAAATGCAAAATTTACAGAGGCTGTGTGTGTGTGTGTGTGTGTGT 1996  
DB 10270 CCGCGTGTATCAAAATGCAAAATTTACAGAGGCTGTGTGTGTGTGTGTGTGTGTGT 10329  
QY 1997 AGTACTTGGAGAGGCTGAGGACAGAGATGCTTTGAACCCAGAGAGTGTGACATG 2056  
DB 10330 AGTACTTGGAGAGGCTGAGGACAGAGATGCTTTGAACCCAGAGAGTGTGACATG 10389  
QY 2057 AGCTGATGTGCTGCTGT 2116  
DB 10390 AGCTGATGTGCTGCTGT 10449  
QY 2117 AAAAAAAG 2176  
DB 10450 AAAAAAAG 10509  
QY 2177 AAGCATGCAAGAAATGTGCTTTTCCAAAGAAATGTAAGAGCAATTAAGAGATCC 2236  
DB 10510 AAGCATGCAAGAAATGTGCTTTTCCAAAGAAATGTAAGAGCAATTAAGAGATCC 10569  
QY 2237 TATTTGCTCTTTTGT 2296  
DB 10570 TATTTGCTCTTTTGT 10629  
QY 2297 TTCAATATCAATATCCCTGT 2356  
DB 10630 TTCAATATCAATATCCCTGT 10689  
QY 2357 AGATCCACAGAGAAACCTGATGACAACTGTCTTATTTTAAATCTTATGTCATTAAGTT 2416  
DB 10690 AGATCCACAGAGAAACCTGATGACAACTGTCTTATTTTAAATCTTATGTCATTAAGTT 10749  
QY 2417 TGTAAAAAGATTAATAATGTTACTCATGTATTAATTAATTAATTAATTAATTAAT 2476

[illegible]

Rives,M., Rojes,A., Rojuboakan,I., Rolfe,N., Ruiz,S., Savery,G., Scherer,E., Scott,G., Shen,H., Shim,C., Shooshitari,N., Sisson,I., Sodergren,J., Sonalike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umami,K., Vazquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,U., Zorrilla,S., Kucherlapati,R., Meinstock,G. and Gibbs,R.

TITLE  
JOURNAL  
Unpublished

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (25-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (25-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT  
On May 25, 2002 this sequence version replaced gi:20901754.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:  
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.  
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.  
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarly (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://hgsc.bcm.tmc.edu:8068/quality/info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
Source	1. .71132 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="12" /clone="RP11-438L7" /note="overlaps bases 1. .1448 of clone AC092490" /function="clone overlap"
misc_feature	187. .408 /standard_name="57233"
STS	439. .560 /standard_name="92005"
STS	774. .881 /rpt_family="MIR"
repeat_region	903. .1190 /rpt_family="AluSg"
repeat_region	1191. .1213 /rpt_family="AT_rich"
STS	1591. .1807 /standard_name="6612"
STS	1744. .1819 /standard_name="8198"
repeat_region	1966. .2264 /rpt_family="AluSg"
repeat_region	3296. .3328 /rpt_family="(TTTC)n"
repeat_region	4087. .4249 /rpt_family="AluSg"
repeat_region	4360. .4654 /rpt_family="AluY"
repeat_region	complement(5140. .5262) /rpt_family="FLAM_C"
repeat_region	5275. .5304 /rpt_family="AT_rich"
repeat_region	5305. .5563 /rpt_family="AluDo"
repeat_region	5564. .5589 /rpt_family="AT_rich"
repeat_region	5644. .5733 /rpt_family="MSTD"
repeat_region	5734. .6012 /rpt_family="AluSg"
repeat_region	6020. .6336 /rpt_family="AluSg"
repeat_region	6337. .6629 /rpt_family="MSTD"
repeat_region	6630. .6650 /rpt_family="(CAA)n"
repeat_region	complement(6711. .6817) /rpt_family="MIR"
repeat_region	complement(7033. .7327) /rpt_family="AluSg"
repeat_region	7607. .7745 /rpt_family="FLAM_A"
repeat_region	7980. .8047 /rpt_family="MIR"
repeat_region	complement(8050. .8549) /rpt_family="LTR47A"
repeat_region	8567. .8698 /rpt_family="MIR"
repeat_region	complement(8815. .9101) /rpt_family="AluSg"
repeat_region	complement(11227. .12175) /rpt_family="LTR5"
STS	12625. .12900 /standard_name="87432"
repeat_region	complement(13547. .13664) /rpt_family="LIMB5"
repeat_region	complement(13665. .13979)

Query Match	77.2%; Score 2174.6; DB 9; Length 71132;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2177; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
QY	617 CTTTGGCCCTGTATGAGTTGATGACTTACGAGCGCATTTGCTACTTTGGACTTTGA 676 
DB	43493 CTACAGCCCTGTATGAGTTGATGACTTACGAGCGCATTTGCTACTTTGGACTTTGA 43552 
QY	677 TAGCACTTCAGGAATGTCAACACGATGAATATCTCTGCTGAAGACAGTGATPAA 736 
DB	43553 TAGCACTTCAGGAATGTCAACACGATGAATATCTCTGCTGAAGACAGTGATPAA 43612 
QY	737 AACAGTCTTCAAGTCTCTCTGTTTATTTCTTCAACTCTCACTTTCTTAAGTTTCA 796 
DB	43613 AACAGTCTTCAAGTCTCTCTGTTTATTTCTTCAACTCTCACTTTCTTAAGTTTCA 43672 
QY	797 GAAAAATTTTATATATGAGCTCTTAAAGATCTATGCTTGAATPAGAGAGAAC 856 
DB	43673 GAAAAATTTTATATATGAGCTCTTAAAGATCTATGCTTGAATPAGAGAGAAC 43732 
QY	857 ACAGTCTGCGACGAGACGCTGCTCAATTGTCAGTTTGAATGACATTTGCCCTA 916 
DB	43733 ACAGTCTGCGACGAGACGCTGCTCAATTGTCAGTTTGAATGACATTTGCCCTA 43792 
QY	917 CTGGGAATPACAGACCTGCGACGCTGGAGCATCTTAAAGTGTCAAGTTTCTATGA 976 
DB	43793 CTGGGAATPACAGACCTGCGACGCTGGAGCATCTTAAAGTGTCAAGTTTCTATGA 43852 
QY	977 CTTTATGATGAGAGAGAGAGAGATCTTAAAGATCTTAAAGATGAGAGATGCAAG 1036 
DB	43853 CTTTATGATGAGAGAGAGAGATCTTAAAGATCTTAAAGATGAGAGATGCAAG 43912 
QY	1037 TTTTATATCAACATCTTATTTATTTGATTTCAATTTGATTAACAGTGTGATGATA 1096 
DB	43913 TTTTATATCAACATCTTATTTATTTGATTTCAATTTGATTAACAGTGTGATGATA 43972 
QY	1097 GATTTTCTATTTCTTTCCCTTGAAGTTTCACTTCAATPACAACTCTTCATCAGG 1156 
DB	43973 GATTTTCTATTTCTTTCCCTTGAAGTTTCACTTCAATPACAACTCTTCATCAGG 44032 
QY	1157 CCATGATCTATAGACCTCTTATGAGATCTGAGTATTTGATGCCCAACCATCTC 1216 
DB	44033 CCATGATCTATAGACCTCTTATGAGATCTGAGTATTTGATGCCCAACCATCTC 44092 
QY	1217 TCAGAGATTAATATATCAATGAGCTGATGTTTATATGACAGAGACATGTTTGA 1276 
DB	44093 TCAGAGATTAATATATCAATGAGCTGATGTTTATATGACAGAGACATGTTTGA 44152 
QY	1277 TGTTTGACAAAGAGATTGTATGAGTGGAGATGAGAGTATAGACATGATGTGAC 1336 
DB	44153 TGTTTGACAAAGAGATTGTATGAGTGGAGATGAGAGTATAGACATGATGTGAC 44212 
QY	1337 CTTCAAGTACTTAAATTAAGATCTTAAATGAGGACGAGTGTGAACAGACACCC 1396 
DB	44213 CTTCAAGTACTTAAATTAAGATCTTAAATGAGGACGAGTGTGAACAGACACCC 44272 
QY	1397 TAAATATGAGTGTATGCTGAGTGAAGCAATCTTGGAAAGCAACTTTTAAAGAA 1456 
DB	44273 TAAATATGAGTGTATGCTGAGTGAAGCAATCTTGGAAAGCAACTTTTAAAGAA 44332 
QY	1457 GTCCCTAATTAGAAACCCACAACTTCAATATCATATTAGCAACATTTGAGAGG 1516 

Db	44333	GTCCCTAATTAGAAACCAACCAACTTCACATATCATATTAAGCAACATTTGGAAG	44332
OY	1517	AAGTTGCTTGAAATGTTGGGAGAGAAATTTATTTGCTCTCGTGGGTCTTTCATCTCA	1576
Db	44393	AAGTTGCTTGAAATGTTGGGAGAGAAATTTATTTGCTCTCGTGGGTCTTTCATCTCA	44452
OY	1577	GAAATGCATACAGGTCAGAGTTTGCAATTTGTATGTGTGAATGCTTCCCAAG	1636
Db	44453	GAAATGCATACAGGTCAGAGTTTGCTCATTTTGTATGTGTGAATGCTTCCCAAG	44512
OY	1637	GTATATTACTATATAGAGATTTGTGCAAAACAGATGATTAAGTGGCAACGTGGC	1696
Db	44513	GTATATTACTATATAGAGATTTGTGCAAAACAGATGATTAAGTGGCAACGTGGC	44572
OY	1697	ACAAGCTCATAGTTCTAGCTCTTTGGAGGTTGAGAGGAGATGGCTTGAACACAGGT	1756
Db	44573	ACAAGCTCATAGTTCTAGCTCTTTGGAGGTTGAGAGGAGATGGCTTGAACACAGGT	44632
OY	1757	GTTCAAGGCGAGCCGTGGGCAACATTAACAATCTGTCTCTCAAAAAAAAAAAAAA	1816
Db	44633	GTTCAAGGCGAGCCGTGGGCAACATTAACAAGTCTGTCTCTCAAAAAAAAAAAAAA	44692
OY	1817	AAGAAAGAGAGAGGGCCGGCGTGTGGTCAACGCTGTATCCCAAGCTTTGGAGGGC	1876
Db	44693	AAGAAAGAGAGAGGGCCGGCGTGTGGTCAACGCTGTATCCCAAGCTTTGGAGGGC	44752
OY	1877	CGAGCCGGGCGGATCACTGTGTGCAGAGATTTGAGCAAGCCTGGCAACATGGCAAA	1936
Db	44753	CGAGCCGGGCGGATCACTGTGTGCAGAGATTTGAGCAAGCCTGGCAACATGGCAAA	44812
OY	1937	CCCCGTCTGTACTCAAAATGCAAAATTTAGCCAGGCGTGTAGCAGACCTGTATCCC	1996
Db	44813	CCCCGTCTGTACTCAAAATGCAAAATTTAGCCAGGCGTGTAGCAGACCTGTATCCC	44872
OY	1997	AGCTACTTGGAGGCTGAGCAGAGAAATGCTTGAACCCAGAGGTGAGAGTTGCAAT	2056
Db	44873	AGCTACTTGGAGGCTGAGCAGAGAAATGCTTGAACCCAGAGGTGAGAGTTGCAAT	44932
OY	2057	AGCTGAGATCGTGGCGTTGCACTCCAGCCTGGGCGAACAGGCMAACTCTGCTCAGAA	2116
Db	44933	AGCTGAGATCGTGGCGTTGCACTCCAGCCTGGGCGAACAGGCMAACTCTGCTCAGAA	44992
OY	2117	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAAAGAACATATTTGGAGAGAAAGATGGG	2176
Db	44993	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAAAGAACATATTTGGAGAGAAAGATGGG	45052
OY	2177	AAGCATTTGCAAGGAAATTTGCTTTATCCAAACAAATGTATAGAGGCCAATAGGATCCC	2236
Db	45053	AAGCATTTGCAAGGAAATTTGCTTTATCCAAACAAATGTATAGAGGCCAATAGGATCCC	45112
OY	2237	TATTTGCTCTTTTGATGTCTATTTTGCCTTAACAATGCTCTTTGACATGAGAAAAAT	2296
Db	45113	TATTTGCTCTTTTGATGTCTATTTTGCCTTAACAATGCTCTTTGACATGAGAAAAAT	45172
OY	2297	TTCAAGATTAACATATCCCTGTGCCATTATTAACCTTGCAATGAAGATGAGC	2356
Db	45173	TTCAAGATTAACATATCCCTGTGCCATTATTAACCTTGCAATGAAGATGAGC	45232
OY	2357	AGATCCACAGGAAACCTTGAATGCAACATGCTTTATTTTATCTTATTTGTACATAGTT	2416
Db	45233	AGATCCACAGGAAACCTTGAATGCAACATGCTTTATTTTATCTTATTTGTACATAGTT	45292
OY	2417	TGTAAAAAGATTTAAAAATTTGTACTCATGATATTCATTAATTTTAAATTAATTTTGGGT	2476
Db	45293	TGTAAAAAGATTTAAAAATTTGTACTCATGATATTCATTAATTTTAAATTAATTTTGGGT	45352
OY	2477	CTAATGATTTTATTAATCAATGATTTCTTTTCTGATATTAATTTGAATGAGATCTCAAGC	2536
Db	45353	CTAATGATTTTATTAATCAATGATTTCTTTTCTGATATTAATTTGAATGAGATCTCAAGC	45412
OY	2537	TTCAATTAATTTAATCTTTAGAAATGATTTCTAATTAACAAGATGATTTGTAACATTTGC	2596

[illegible]



OY	863	CTGGCAGAGGAGCGTGTCTCAATGGTGTGAGTTTGAAGCAACATGTGCCCCCTACTGGGA	922
Db	241	CTGGCAGAGGAGCGTGTCTCAATGGTGTGAGTTTGAAGCAACATGTGCCCCCTACTGGGA	300
OY	923	ATAACAGAACTGCAGAGACCTGGAGCAATCCATAAGTGTCAACGTTTTCTATGACTTTTA	982
Db	301	ATAACAGAACTGCAGAGACCTGGAGCAATCCATAAGTGTCAACGTTTTCTATGACTTTTA	360
OY	983	GGTAGGATGAGAGCAGAAAGTATGATCTTAAAAAGCATGTGAGAGGATCAAAATGTTTTTA	1042
Db	361	GGTAGGATGAGAGCAGAAAGTATGATCTTAAAAAGCATGTGAGAGGATCAAAATGTTTTTA	420
OY	1043	TATCAACATCCCTTATTTATTTGATTCATTTAGTTTAAAGTGTGTGTGTATGATTTT	1102
Db	421	TATCAACATCCCTTATTTATTTGATTCATTTGATTTAAAGTGTGTGTGTATGATTTT	480
OY	1103	TCATATCTTTTCCCTTGACGTTTACTTCAAGTAAACAACAATCTTCATCAGGCGATGA	1162
Db	481	TCATATCTTTTCCCTTGACGTTTACTTCAAGTAAACAACAATCTTCATCAGGCGATGA	540
OY	1163	TCATATGAGACCTCCTTAATGAGATATCTGGGTGATTGTACCCCAACCATCTCTCCAA	1222
Db	541	TCATATGAGACCTCCTTAATGAGATATCTGGGTGATTGTACCCCAACCATCTCTCCAA	600
OY	1223	GCATTAATATCCATCATATGCGCTGTATGTTTTAATCAGCAGAAAGATGTTTTATGTTTG	1282
Db	601	GCATTAATATCCATCATATGCGCTGTATGTTTTAATCAGCAGAAAGATGTTTTATGTTTG	660
OY	1283	TACAAAAAAGATGTTTATGAGTGGGAGATGAGAGATATGACATGATGATCAGCTTCAA	1342
Db	661	TACAAAAAAGATGTTTATGAGTGGGAGATGAGAGATATGACATGATGATCAGCTTCAA	720
OY	1343	GCTACTTTAATTAAGATCTTAAATGAGGAGAGAGACTGTGAACAAGACACCTTAATTA	1402
Db	721	GCTACTTTAATTAAGATCTTAAATGAGGAGAGAGACTGTGAACAAGACACCTTAATTA	780
OY	1403	TGGGTGATGTCTGAGTAGCAAAATCTTCTGAAAACGCAAACTCTTTTAAGAAAGTCCCT	1462
Db	781	TGGGTGATGTCTGAGTAGCAAAATCTTCTGAAAACGCAAACTCTTTTAAGAAAGTCCCT	840
OY	1463	AATTTAGAAACACCCACAACTTTCACATATCATTAATTGCAAAACAATTGGAAGAAAGTTG	1522
Db	841	AATTTAGAAACACCCACAACTTTCACATATCATTAATTGCAAAACAATTGGAAGAAAGTTG	900
OY	1523	CTTGAATTTGGGAGAGAAATCTATTTGGCTCTCGGGGCTCTTCATCTCAGAAATG	1582
Db	901	CTTGAATTTGGGAGAGAAATCTATTTGGCTCTCGGGGCTCTTCATCTCAGAAATG	960
OY	1583	CCAATCAGGTCAAGGTTTGCTACATTTTGTATGTGTGTATGATGCTTCTCCAAAGGTATAT	1642
Db	961	CCAATCAGGTCAAGGTTTGCTACATTTTGTATGTGTGTATGATGCTTCTCCAAAGGTATAT	1020
OY	1643	TAACTATTTAAGAGATGTGTGACAAAACAGATATATAAGTCTGCAACCGTGTGACACGC	1702
Db	1021	TAACTATTTAAGAGATGTGTGACAAAACAGATATATAAGTCTGCAACCGTGTGACACGC	1080
OY	1703	TCATATGTTCTTACTGCTCTGGAGAGTTTGAGAGAGAGATGAGCTTGTAACAAGTGTCTAA	1762
Db	1081	TCATATGTTCTTACTGCTCTGGAGAGTTTGAGAGAGAGATGAGCTTGTAACAAGTGTCTAA	1140
OY	1763	GGCCAGCTCGGGCAACATTAACAGATCCTGTCTCTCAAAAAAAAAAAAAAAAAAAGAA	1822
Db	1141	GGCCAGCTCGGGCAACATTAACAGATCCTGTCTCTCTCAAAAAAAAAAAAAAAAAAAGAA	1200
OY	1823	GAGAGAGGCGCGGGCTGTGTGCTCAACGCTGTATATCCAGACATTTGGAGGCGGAGCC	1882
Db	1201	GAGAGAGGCGCGGGCTGTGTGCTCAACGCTGTATATCCAGACATTTGGAGGCGGAGCC	1260
OY	1883	GGGCGGATCAACTGTGTGTCAAGAGATTGTAAGACAGCTCGGCAACATGTGGCAAAACCCGT	1942
Db	1261	GGGCGGATCAACTGTGTGTCAAGAGATTGTAAGACAGCTCGGCAACATGTGGCAAAACCCGT	1320

Oy	1943	TTGAACTCAAAATGCAAAAATTAGCCAGGGGTGGTAGCAGCACCTGTAATCCACGGTAC	2002
Dp	1321	CTGACTCAAAATGCAAAAATTAGCCAGGGGTGGTAGCAGCACCTGTAATCCACGGTAC	1380
Oy	2003	TTGGAGGCTGAGCGAGAGAAATGCTTTGAACCCAGAGGGTGGAGTTTGCAGTAACTGA	2062
Dp	1381	TTGGAGGCTGAGCGAGAGAAATGCTTTGAACCCAGAGGGTGGAGTTTGCAGTAACTGA	1440
Oy	2063	GATGTGCGCTGTGCACTTCACGCTCGGCGCAAGAGCAAGACTGTGCTTCAGAAAAAAA	2122
Dp	1441	GATGTGCGCTGTGCACTTCACGCTCGGCGCAAGAGCAAGACTGTGCTTCAGAAAAAAA	1500
Oy	2123	AAAAAAAAAGAGAGAGAGAAAGAAACAATTTTGGAGAGAAAGATGGGGAAACAT	2182
Dp	1501	AAAAAAAAAGAGAGAGAGAAAGAAACAATTTTGGAGAGAAAGATGGGGAAACAT	1560
Oy	2183	TGCAGAGAAATTTGCTTATTTCCAACAAAATGTATGAGAGCAATTAAGGATTTCCATTTTG	2242
Dp	1561	TGCAGAGAAATTTGCTTATTTCCAACAAAATGTATGAGAGCAATTAAGGATTTCCATTTTG	1620
Oy	2243	TCTCTTTGGTGTCTATTTTCCCTAACAACTGTCTTTGAACAGTGAAGAAAAATTTTCA	2302
Dp	1621	TCTCTTTGGTGTCTATTTTCCCTAACAACTGTCTTTGAACAGTGAAGAAAAATTTTCA	1680
Oy	2303	ATAACCAATTCCTGTGCGCTTATTAACCTTGCAACCTTGCAATGAAGATGAGCAATTC	2362
Dp	1681	ATAACCAATTCCTGTGCGCTTATTAACCTTGCAACCTTGCAATGAAGATGAGCAATTC	1740
Oy	2363	ACAGAGAACTTGAAATGCAACAATGCTTTATTTTAACCTTAATGTATAGTATTGTA	2422
Dp	1741	ACAGAGAACTTGAAATGCAACAATGCTTTATTTTAACCTTAATGTATAGTATTGTA	1800
Oy	2423	AGAGTTAAAAATGTACTTCATGATTCATTTAATTTATATTTATTTTGGCGCTAATG	2482
Dp	1801	AGAGTTAAAAATGTACTTCATGATTCATTTAATTTATATTTATTTTGGCGCTAATG	1860
Oy	2483	ATTTTATTATTAACATGATTTCTTTTCTGATATATTTGAAATGGAAGTCCAAAGCTTCATA	2542
Dp	1861	ATTTTATTATTAACATGATTTCTTTTCTGATATATTTGAAATGGAAGTCCAAAGCTTCATA	1920
Oy	2543	AATTTATTAACCTTTTGAATGTCTTAATAACAACGTATGTAAATTTGTAACTTGCAGTAAT	2602
Dp	1921	AATTTATTAACCTTTTGAATGTCTTAATAACAACGTATGTAAATTTGTAACTTGCAGTAAT	1980
Oy	2603	GGTGTCTGAGAGCAATTCCTGATATTTTGAATGAATCTTTATGACACGCAAAATTTGCTTC	2662
Dp	1981	GGTGTCTGAGAGCAATTCCTGATATTTTGAATGAATCTTTTATGACACGCAAAATTTGCTTC	2040
Oy	2663	TGGCTCACTTTCATCATCGTTAAATTAATTAATAATTTTGGAGCTGTGAAGATTA	2722
Dp	2041	TGGCTCACTTTCATCATCGTTAAATTAATTAATAATTTTGGAGCTGTGAAGATTA	2100
Oy	2723	TACCAATTAATATATATTAATAAGTATTTATATGACGTAAATTAATAATTAATTAATTAAT	2782
Dp	2101	TACCAATTAATATATATTAATAAGTATTTATATGACGTAAATTAATAATTAATTAATTAAT	2160
Oy	2783	GGAATTAACCTTG	2794
Dp	2161	GGAATTAACCTTG	2172

RESULT 8	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BC006296	BC006296	1828 bp mRNA linear	BC006296	BC006296	Hom sapiens activation-induced cytidine deaminase, mRNA (cDNA clone MGC:12911 IMAGE:4054915), complete cds.	GI:3871601	Hom sapiens (human)
							Hom sapiens
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 1828)  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schlier G.D.,  
Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,  
Scheetz T.E., Brownstein M.J., Ueda T.B., Toshiyuki S.,  
Garcia J.P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,  
Abramson R.D., Muliyil S.J., Bosak S.A., McEwan P.J.,  
McKernan K.J., Malek J.A., Gamaralline P.H., Richards S.,  
Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S.,  
Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,  
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,  
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S., Krzywinski M.I., Skalske U., Smallus D.E.,  
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
2 (bases 1 to 1828)  
Strausberg R.  
Direct Submission  
Submitted (09-APR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
On Aug 19, 2003 this sequence version replaced gi:13623400.  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Louis Straud  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Achter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,  
Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,  
Dietrich N.L., Granite S., Guan X., Gupta J., Haghghi P.,  
Hansen N., Ho S.-L., Karlins E., Kwong P., Lurie P., Legaspi R.,  
Maduro Q.L., Mastello C., Maskeri B., Mastrlian S.D., McCloskey J.C.,  
McDowell J.J., Pearson R., Stantipop S., Thomas P.J., Touchman J.W.,  
Tsurgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,  
Young A., Zhang L.-H. and Green E.D.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/URL at: <http://image.llnl.gov>  
Series: IRAL Plate: 17 Row: a Column: 1  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 10190699.  
Location/Qualifiers  
1. 1828  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:12911 IMAGE:4054915"  
/cissue\_type="Primary B-Cells from Tonsils"  
/clone\_id="NIH\_MGC\_48"  
/lab\_host="DH10B-R"  
/note="Vector: pOT87"  
1. 1828  
/gene="AICDA"  
/note="synonyms: AID, HIGM2, CDA2, ARP2"  
/db\_xref="Locustid:57379"  
/db\_xref="MIM:605257"  
77. 673

ORIGIN  
Query Match 64.5%; Score 1818.6; DB 9; Length 1828;  
Beet Local Similarity 99.8%; Pred. No. 0;  
Matches 1821; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
/codon\_start=1  
/product="activation-induced cytidine deaminase"  
/protein\_id="AAH06296.1"  
/db\_xref="GI:13623401"  
/db\_xref="locustid:57379"  
/transcript="M5SLMNRRELYOPFNWAKRRELYCYVWRSDSPTSPST  
DFGLRNKNGCHELLELRISPDUDPGRCYRTWTMSPTCDKRPADLRNRP  
NLSRLITFARLFCEDKRPEGRRLRGRGVAIATGFDYFCWNTFVENHERFEK  
AMEGLHNSVRLSRQLRIHLLPEYEDVDLDRFRTGL"  
4 GAACCATTAATTAATGAAGTAGAATTTTCTGCGCTGAGACTTGACAGGAGCAAGAGA 63  
1 GAACCATTAATTAATGAAGTAGAATTTTCTGCGCTGAGACTTGACAGGAGCAAGAGA 60  
64 CACTTGAGACAACATAGACAGCCTTGATGATGACCGAGAGAAATTTCTTACCAATT 123  
61 CACTTGAGACAACATAGACAGCCTTGATGATGACCGAGAGAAATTTCTTACCAATT 120  
124 CAAAATGTCGCTGCGCTGAGAGGCTGAGAGCTGAGAGCTTACTGCTAGCTAGTAAAG 183  
121 CAAAATGTCGCTGCGCTGAGAGGCTGAGAGCTTACTGCTAGCTAGTAAAG 180  
184 GCGTGAAGAGTCTACATCCTTTTCTGAGAGCTTGTGTTATCTTGCAATAGAAAGGCTG 243  
181 GCGTGAAGAGTCTACATCCTTTTCTGAGAGCTTGTGTTATCTTGCAATAGAAAGGCTG 240  
244 CCAAGTGAATGCTCTTCTCCGCTACATCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTG 303  
241 CCAAGTGAATGCTCTTCTCCGCTACATCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTG 300  
304 CTACCGGCTCACCTGCTTACCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 363  
301 CTACCGGCTCACCTGCTTACCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 360  
364 CGACTTCTGAGAGAGAAACCCCAACCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 423  
361 CGACTTCTGAGAGAGAAACCCCAACCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 420  
424 CTGAG 483  
421 CTGAG 480  
484 AATGAGCATATGACCTTCAAGATTATTTTCTGAGAGCTGAGAGCTGAGAGCTGAGAG 543  
481 AATGAGCATATGACCTTCAAGATTATTTTCTGAGAGCTGAGAGCTGAGAGCTGAGAG 540  
544 TGAAGAAGCTTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603  
541 TGAAGAAGCTTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
604 GCTTGGGAGATCTTTTGGCCCTGATGAGAGTGAATTTTCAAGAGAGAGAGAGAGAG 663  
601 GCTTGGGAGATCTTTTGGCCCTGATGAGAGTGAATTTTCAAGAGAGAGAGAGAGAG 660  
664 TTTGGAGCTTGAATGAGACCTTCAAGAGTGAATTTTCAAGAGAGAGAGAGAGAGAG 723  
661 TTTGGAGCTTGAATGAGACCTTCAAGAGTGAATTTTCAAGAGAGAGAGAGAGAGAG 720  
724 ACAATGATTAATAAAGAGCTCTTCAAGTCTTCTGTTTATTTTATTTTCAACTCTT 783  
721 ACAATGATTAATAAAGAGCTCTTCAAGTCTTCTGTTTATTTTATTTTCAACTCTT 780  
784 CTTAGATTTTAAAGAAAAATTTTATATACAGACTCTTTTAAAAAGATCTATGCTTGA 843  
781 CTTAGATTTTAAAGAAAAATTTTATATACAGACTCTTTTAAAAAGATCTATGCTTGA 840  
844 ATAG 903

Db 841 ATAGAGAGAGACAGGTCTGGCAGGAGCTGCTGCAATGTCAGATTTTGAATGCA 900  
Qy 904 ACATTTGCCCTACTGGGAATTAACAGACCTGCGAGCTGGAGCATCTCTAAAGTCTCA 963  
Db 901 ACATTTGCCCTACTGGGAATTAACAGACCTGCGAGCTGGAGCATCTCTAAAGTCTCA 960  
Qy 964 CGTTTTCATGACTTTAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1023  
Db 961 CGTTTTCATGACTTTAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
Qy 1024 AGAGATCAAAATGTTTATATCAACATCTCTATTTATTTGATTCATTTAGTTAAAGT 1083  
Db 1021 AGAGATCAAAATGTTTATATCAACATCTCTATTTATTTGATTCATTTAGTTAAAGT 1080  
Qy 1084 GGTGTAGTATGATTTTCTATTTCTTTTCCCTTGAAGTTTCACTTTCAAGTAAACAAA 1143  
Db 1081 GGTGTAGTATGATTTTCTATTTCTTTTCCCTTGAAGTTTCACTTTCAAGTAAACAAA 1140  
Qy 1144 CTCTTCATGAGGCGATGATCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1203  
Db 1141 CTCTTCATGAGGCGATGATCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
Qy 1204 CCCAAACATCTCTCCAAAGCATTAATATCAATCATGCGCTGTATGTTTATATGAGAG 1263  
Db 1201 CCCAAACATCTCTCCAAAGCATTAATATCAATCATGCGCTGTATGTTTATATGAGAG 1260  
Qy 1264 AAGCATGTTTATGTTTGTACAAAAGAGATTTGATGAGGAGTGAAGTATAGAC 1323  
Db 1261 AAGCATGTTTATGTTTGTACAAAAGAGATTTGATGAGGAGTGAAGTATAGAC 1320  
Qy 1324 CAGCATGTCACCTTCAAGCTCTTATTAAGAGATCTTAAATGAGAGAGAGAGAGAGAG 1383  
Db 1321 CAGCATGTCACCTTCAAGCTCTTATTAAGAGATCTTAAATGAGAGAGAGAGAGAGAG 1380  
Qy 1384 GAACAAGACACCTTAATATGAGTGAATGTCAGAGTGAAGTCTTCCGAGAGAGAGAGAG 1443  
Db 1381 GAACAAGACACCTTAATATGAGTGAATGTCAGAGTGAAGTCTTCCGAGAGAGAGAGAG 1440  
Qy 1444 CTCTTTTAAAGAGTCCCTAATTTAGAAACACCAAACTTCAATATCATATTTAGCA 1503  
Db 1441 CTCTTTTAAAGAGTCCCTAATTTAGAAACACCAAACTTCAATATCATATTTAGCA 1500  
Qy 1504 AACATTTGAGAGAGTGTGTTGAATGTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1563  
Db 1501 AACATTTGAGAGAGTGTGTTGAATGTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560  
Qy 1564 TCTCTTCATCTCGAATATGCAATCATGCTCAAGTTCATCATTTTGTATGTGTGAT 1623  
Db 1561 TCTCTTCATCTCGAATATGCAATCATGCTCAAGTTCATCATTTTGTATGTGTGAT 1620  
Qy 1624 GCTTCTCCCAAGGTATTAATTAATATTAAGAGAGTGTGACAAAACGAATGATTAAGC 1683  
Db 1621 GCTTCTCCCAAGGTATTAATTAATTAAGAGAGTGTGACAAAACGAATGATTAAGC 1680  
Qy 1684 TGGGAACCGTGACACGCTCATAGTCTAGCTGCTTGGAGAGTGTGAGAGAGAGAGAGAG 1743  
Db 1681 TGGGAACCGTGACACGCTCATAGTCTAGCTGCTTGGAGAGTGTGAGAGAGAGAGAGAG 1740  
Qy 1744 CTGGAACACAGAGTGTCAAGAGCGCTGGGCAATTAACAGATCTCTCTCAAAA 1803  
Db 1741 CTGGAACACAGAGTGTCAAGAGCGCTGGGCAATTAACAGATCTCTCTCAAAA 1800  
Qy 1804 AAAAAAAAAAAAAAAAAAGAGAGA 1828  
Db 1801 AAAAAAAAAAAAAAAAAA 1825

RESULT 9  
BD016828 2440 bp DNA linear PAT 27-AUG-2002  
LOCUS Novel cytidine deaminase.  
DEFINITION  
ACCESSION BD016828  
VERSION BD016828.1 GI:22558004

KEYWORDS JP 2001245669-A/1.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS Honjo, T. and Muramatsu, M.  
TITLE Novel cytidine deaminase  
JOURNAL Patent: JP 2001245669-A 1 11-SEP-2001;  
JAPAN TOBACCO INC, TASUKU HONDO  
OS Mus musculus (mouse)  
PN JP 2001245669-A/1  
PD 11-SEP-2001  
PF 28-MAR-2000 JP 2000092981  
PI TASUKU HONDO, NASAMICHI MURAMATSU  
PC C12N15/09, A61K39/395, A61K39/395, A61P1/00, A61P1/06, A61P13/12,  
PC A61P17/00,  
PC A61P27/02, A61P37/16, A61P37/02, A61P37/08, C07K16/18, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12N5/78, C12P21/02, C12P21/08, C12N1/21, C12N1/19, PC  
(C12N5/10, C12N1/91), C12N15/00, C12N5/00, (C12N5/00, C12N1/91) CC  
FH Key Location/Qualifiers  
FT CDS (93) (689)  
FT 5'UTR (1) (92)  
FT 3'UTR (690) (2440).

FEATURES  
source 1..2440  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"

ORIGIN  
Query Match 21.4%; Score 603.4; DB 6; Length 2440;  
Best Local Similarity 69.4%; Pred. No. 6,4e-108;  
Matches 1006; Conservative 0; Mismatches 351; Indels 92; Gaps 10;  
Qy 4 GAACCATCAATTAATGAAGTGAATTTTCTGAGCTGAGACTTGAAGAGGAGAG-AG 62  
Db 16 GAAGACACCTTGCTTGAAGCAGACTTCTTGAGCTTGAAGCTTGAAGAGTCAAGAAAG 75  
Qy 63 ACATCTGAGACACCATATGAGACAGCTCTGTATGAACGAGAGAGAGTTCTTTACCAAT 122  
Db 76 TCACGCTGAGAGACCATATGAGACAGCTCTGTATGAAGAGAGAGAGTTCTTTACCAAT 135  
Qy 123 TCAAAATATGCGCTGGGCTAAGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 182  
Db 136 TCAAAATATGCGCTGGGCTAAGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 195  
Qy 183 GAGCTGACAGTGTACATCTCTTCACTGAGACTTGTATCTTCCCAATTAAGAGCGCT 242  
Db 196 GAGAGATATGCACTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 255  
Qy 243 GCAAGTGAATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302  
Db 256 GCAAGTGAATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 315  
Qy 303 GCAAGTGAATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362  
Db 316 GCAAGTGAATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 375  
Qy 363 CCGACTTCTGAG 422  
Db 376 CCGACTTCTGAG 435  
Qy 423 TCTGTGAGAGCGCAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482  
Db 436 TCTGTGAGAGCGCAAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 495  
Qy 483 AATATGCAATCAAG 542  
Db 496 AATATGCAATCAAG 555  
Qy 543 ATGAAGACATTTCAAG 602

```

Db      556 GTAAAGAACTTCAAGCTGGGAAGGGGTACATAAAATTCGTCCGGTACCAAGC 615
Qy      603 AGTTGGGGGACATCCTTTTGGCCCTGTATGAGTTATGACTTACAGACGATTTTCGA 662
Db      616 AACTTGGGGGACATCCTTTTGGCCCTGTATGAGTTATGACTTACAGACGATTTTCGA 675
Qy      663 CTTCGGGACTTGTATGCACTTCCAGGAATGTCAACAGATGAAATATCTCTGTGA 722
Db      676 TGTGGGATTTTAAAGCACTCTCTGAAATGTCAAGTAAATTTCT---CTGA 732
Qy      723 GACAGTGAATPAAAAACAGTCTTCAAGTCTCTCTGTTTATTTCTTCACTCACTT 782
Db      733 GAACTGGAATGAAAAAACAACCTTCA--ACTACAGTTTCTTCTTAAAGTACACTT 790
Qy      783 TCTTAGAGTTTACAGAAAAAATTTATATACAGACTTTTAAAAAGATTTATGCTTGA 842
Db      791 TTATAGTGTAGGGGAAA-----TTATAGACTTTTAAAAAATCTTGAGCTGCAC 843
Qy      843 AATAGAAAGGAACAAGTCTGGCCAGGAGCTGCTCAATTGGTCAGTTTGAATGC 902
Db      844 AGGACCGGCAAGACATGATGTAAC-----TGAGCTTGTCTGTC 882
Qy      903 AACATGTCCCTACT--GGGAATPACAGACTGACAGACCTGGAGACATCTAAAGTGC 961
Db      883 AACATGCCATCTACTGGGGAACAGATTAATCTTCCAGACTTTGGTGTGAATGATCTC 942
Qy      962 AACGTTTTCTATGACTTTAGTAGAGTAGAGAGAGAGATCTTAAAGACATGG 1021
Db      943 TTTTCTTCAAGCAT-----GGAAACCATAT 971
Qy      1022 TGAAGAGTAAATGTTTTTATATCAACATCTTATATTTGATTCATTTAGTTACA 1081
Db      972 GGAAGAGACCAACAGTTGTACACCCACCTGTGTCTTGATTCATTTGAATTCGA 1031
Qy      1082 GTGGTGTGTGATGATTTTCTATCTTTTCCCTTGACGTTTACTTCAAGTAAACACA 1141
Db      1032 GGGGATTCAGTGAAGGATTTCTATCTTCCCTTAAAGGCTCACTTCAAG----- 1084
Qy      1142 AACTTTCATCAGGCAATGATCTATAGAGACCTCTAATAGAGATCTGGTGAATGG 1201
Db      1085 ---GTCCTTTTGTGACAGAGTACGGGCTGTCTCAAGTCTGTCTGAGCAAT---- 1137
Qy      1202 ACCCAAACCATCTTCCAAAGCATTATATCCATCATGCGCTGATGTTTAAATCAGC 1261
Db      1138 -CACAAGCATCTCTCAAAAACATTAATATACAGGACATGCTGATGATTTTCA---- 1191
Qy      1262 AGAAGCATGTTTATGTTTGTACAAAAGAAATTTGTATGGGTGGGATGAGGTATAG 1321
Db      1192 -----CTGCCGTGCTGTTTTTCAATTGTATGTGAAGGAGCTTGGGGTGAATTGA 1245
Qy      1322 ACCATGATGATGATCCTTCAAGCTACTTTATATPAAAGATCTTAAATGGGCGAGAGCT 1381
Db      1246 AGAATGCAAGATGCTCTGGGATGATTTCAATTAAGAGATCTTAAATGAGATGAGAGCT 1305
Qy      1382 GTGAACAAGACACCTTAATATATGGTGTGATGATCTGAAGTGAAGCAAAATCTTCTGAAGCA 1441
Db      1306 AGCAAAATCACTCTGAAGATGATGATGAGCTCAAGAGCAAAATCCCTGGAAGACACA 1365
Qy      1442 AACTCTTTT 1450
Db      1366 GACTCTTTT 1374

```

```

RESULT 10
AF132979 2440 bp mRNA linear ROD 05-AUG-1999
LOCUS Mus musculus activation-induced cytidine deaminase (AID) mRNA,
DEFINITION complete cds.
ACCESSION AF132979.1 GI:5305727
VERSION AF132979.1
KEYWORDS Mus musculus (house mouse)

```

```

ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2440)
Muraiatsu,M., Sankaranand,V.S., Anant,S., Sugai,M., Kinoshita,K.,
Davidson,N.O., and Honjo,T.
TITLE Specific expression of activation-induced cytidine deaminase (AID),
a novel member of the RNA-editing deaminase family in germinal
center B cells
JOURNAL J. Biol. Chem. 274 (26), 18470-18476 (1999)
MEDLINE 99303612
PUBMED 10373455
REFERENCE 2 (bases 1 to 2440)
AUTHORS Muraiatsu,M. and Honjo,T.
TITLE Direct Submision
JOURNAL Submitted (04-MAR-1999) Medical Chemistry, Kyoto University, Konoe
Yoshida Sakyo-Ku, Kyoto 606-8501, Japan
FEATURES
source
1..2440
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..2440
/gene="AID"
93..689
/note="APOBEC-1 homologue"
/product="activation-induced cytidine deaminase"
/protein_id="AA041793.1"
/db_xref="GI:5305728"
/translation="MDSLIMKQKPLVFNQVWAKGRHETLYCYVVKRDSATSCSL
DPGLRKSQSGVHFLFRLYSIDMDPQCYVLTFTMSPCYDCARVAFELRNP
NLSIRFATLVCEDEKAEPELRHAGVIGIMTPKDYICNATVEVNERERFK
AMEGHENSRILRLRLPLIYVVDLRDAFRMLGP"

```

## ORIGIN

```

Query Match 21.4%; Score 603.4; DB 10; Length 2440;
Best Local Similarity 69.4%; Pred. No. 6.4e-108;
Matches 1006; Conservative 0; Mismatches 351; Indels 92; Gaps 10;

```

```

Qy      4 GAACCATCATTAATTAAGTGAATTTTCTGCTCTGAGACTTGCAGGAGCAAG-AAG 62
Db      16 GAACCACTCTGCTTGAAGCAAGCTTCTTGCTTAAGACTTTGAGGAGTCAAGAAAG 75
Qy      63 AACTCTGGACACCATATGAGACAGCTCTTATGAAACCGAGAGATTTCTTTACCAAT 122
Db      76 TCAGCTGGAGACCGATATGAGACAGCTTCTGATGAAGCAAAAGATTTCTTTACCAAT 135
Qy      123 TCAAAATATCCGCTGGGCTTAAGGGTGGGCTGAGACCTTCTGTGCTACGTAGTGAAG 182
Db      136 TCAAAATATCCGCTGGGCTTAAGGGTGGGCTGAGACCTTCTGTGCTACGTAGTGAAG 195
Qy      183 GAGCTGACAGTGTACATCTTTTCACTGACCTTTGTTATCTTCCAAATPAAAGCGCT 242
Db      196 GAGAGATATGACACCTCTGCTCACTGAGACTTGGCAACCTTCCAAATPAAAGCGCT 255
Qy      243 GCAACGTGGAATGCTTTCTTCCGCTACATCTGAGACTTGGAGACTTGAACCTTGGCGCT 302
Db      256 GCAACGTGGAATGCTTTCTTCCGCTACATCTGAGACTTGGAGACTTGAACCTTGGCGCT 315
Qy      303 GCTACCGGCTCACTGGTTCACCTCTGAGAGCCCTGCTACAGACTGAGCCGACATGTGG 362
Db      316 GTTACCGGCTCACTGGTTCACCTCTGAGAGCCCTGCTACAGACTGAGCCGACATGTGG 375
Qy      363 CCGACTTTCGAGAGGGAACCCCAACTCAGTGTGAGATCTTCAACCGGCGCTTACT 422
Db      376 CTGAGTTTCTGAGATGGAACCTTACCTCAGCTGAGATTTTCAACCGGCGCTTACT 435
Qy      423 TCTGTGAGAGCGGAAAGCTGAGCCGAGGGGCTGGCGGCTGCAACCGGCGGGGCTGC 482
Db      436 TCTGTGAGAGCGGAAAGCTGAGCCGAGGGGCTGGCGGCTGCAACCGGCGGGGCTGC 495

```

```

QY 483 AATAGCCATCATGACCTTCAAGATTTATTTTACTGCTGGAATCTTTGTGAAGACC 542
Db 496 AGATCGGATCATACCTTCAAGACATATTTTATCTGCTGGAATCTTTGTGAAGACC 555
QY 543 ATGAAGAATCTTCAAGCTGGAAGGCTGACATGAATTCATGCTCTCTCCAGAC 602
Db 556 GTGAAGAATCTTCAAGCTGGAAGGCTGACATGAATTCATGCTCTCTCCAGAC 615
QY 603 AGCTCGGATCTTCTTGGCCCTGTATGAGTTGATGCTTACGAGACGCTTTGCTGA 662
Db 616 AACTTCGGCGATCTTCTTGGCCCTGTATGAGTTGATGCTTACGAGACGCTTTGCTGA 675
QY 663 CTTCGGGACTTGTATGACATCTTCAGAAATGTACACAGATGAATATCTGCTGA 722
Db 676 TGTTCGGATTTTGAAGACATCTCTGGAATGTACAGATGAATATCTGCTGA 732
QY 723 GACAGTGAATGAAGACATCTTCAGATCTCTGCTTATTTTATCTTCACTCTCACTT 782
Db 733 GAGACTGATGAAGACATCTTCAGATCTCTGCTTATTTTATCTTCACTCTCACTT 790
QY 783 TCTTAGATTTACGAAAAATTTATATACGATCTTTTAAAAAGATCTTATGCTTGA 842
Db 791 TTATATAGTTAGGGGAAA-----TTATATGATCTTTTAAAAAATCTTGAAGTGCAC 843
QY 843 AATAGAGAAGAACACAGCTCTGCGCAGGACGCTGCAATTTGTCAGATTTTGAATGC 902
Db 844 AGGACCCGACAGAACATGATGTATAC-----TGACCTTCTGCTGC 882
QY 903 AACATTTGCTCCCTACT--GGGAATPACAGAACTGACAGACCTGGAGACATCTTAAGTGC 961
Db 883 AACATTTGCTCCCTACT--GGGAATPACAGAACTGACAGACCTGGAGACATCTTAAGTGC 942
QY 962 AACGTTTTCTATGACTTTTAGTAGAGTGAAGACAGAGATGATCTTAAAAAGCATGG 1021
Db 943 TTTTTCACACAGAT-----GGAAAAAGCATAT 971
QY 1022 TGAAGAGTCAAAAGTTTATATATCAATCTTATATTTGATTCATTTGATTAACA 1081
Db 972 GGAACAGACACACAGATTTTATATCAATCTTATATTTTATTCATTTGATTAACA 1031
QY 1082 GTGCTGTAGATGATGATTTTCTATCTTCTTCCCTGAGCTTACTTCAAGTAAACACA 1141
Db 1032 GGGGTATCATGATGACGATTTCTTATCTTCTTCCCTGAGCTTACTTCAAGTAAACACA 1084
QY 1142 AACTTTCATCAGGACATGATCTATAGACCTCTTATGAGATATCTGGATGATG 1201
Db 1085 ---GGCTTTTTCAGACAGGTGACGGGGCTGCTTCACTCTGCTGACGAAT----- 1137
QY 1202 ACCCAAAACATCTCTCCAAAGCATTAATATCCAAATCATGCGCTGATGTTTAAATCAG 1261
Db 1138 -CAAAACCAATCTCTCAAAAACATTAATATCTGAGCAATGCTGATATGTTTCA----- 1191
QY 1262 AGAAGCATGTTTATGTTTGTATACAAAAGAAATGTTTATGAGTGGGAGTGAAGTATAG 1321
Db 1192 -----CTGTCGCGTGTGTTTTCATATTTGTATGTAAGAAAGGCTTGGGGTGGAGTTTGA 1245
QY 1322 ACCATGATGATGATCTTCAAGCTTCACTTATATTAAGATCTTAAATAGGGACAGAGACT 1381
Db 1246 AGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
QY 1382 GTGAACAAGACATCTTATATATGAGTATGATGATGATGATGATGATGATGATGATGAT 1441
Db 1306 ACAGAAAGATCACTCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1365
QY 1442 AACTCTTTT 1450
Db 1366 GACTCTTTT 1374

```

```

RESULT 11
LOCUS AF529828 597 bp mRNA linear SYN 17-SEP-2002
DEFINITION Mus musculus clone 1 transgenic Homo sapiens AID (AID) mRNA,

```

```

ACCESSION complete cds.
VERSION AF529828
KEYWORDS AF529828.1 GI:22297243
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL 1 (bases 1 to 597)
TITLE Martin, A. and Schaff, M.D.
REFERENCE Somatic hypermutation of the AID transgene in B and non-B cells
AUTHORS Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
JOURNAL 2 (bases 1 to 597)
TITLE Martin, A. and Schaff, M.D.
REFERENCE Direct Submission
JOURNAL Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
JOURNAL Medicine, 1300 Morris Park Ave. Bronx, NY 10461, USA
FEATURES
source
location/Qualifiers
1..597
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1"
/cell_type="hybridoma P1-5"
/transgenic
1..597
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..597
/gene="AID"
1..597
/gene="AID"
/codon_start=1
/product="AID"
/protein_id="PAM5415.1"
/db_xref="GI:22297244"
/translation="MDSILNRRKFLVQFNVRWAKRRETYLCYVVRKDSATSFEL
DGYLRNNDCHVELFLARYISDWLDPGRCYRTWFTSPSCYCDARHYADFGRNP
NLSRIFFARLYPCEDRKABEGRLRHRAVGQALWTFPDYFYCWNTFENHRTFK
AMBLHNSVRLSRLRLILPLYEVNDLADAPRTGL"
ORIGIN
Query Match 21.1%; Score 595.4; DB 12; Length 597;
Best Local Similarity 99.8%; Pred. No. 2,2e-106;
Matches 596; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 80 ATGACAGCTCTTGATGACACGAGAGAGTTCTTTACCAATCAAAAATGCTCGCTGG 139
Db 1 ATGACAGCTCTTGATGACACGAGAGAGTTCTTTACCAATCAAAAATGCTCGCTGG 60
QY 140 GCTAAGGTCGGCGTGAAGCTAAGCTGCTGCTAAGTGAAGAGGCGTGAAGTCTACA 199
Db 61 GCTAAGGTCGGCGTGAAGCTAAGCTGCTGCTAAGTGAAGAGGCGTGAAGTCTACA 120
QY 200 TCTTTTCACTGACCTTGTATCTTCTGCAATTAAGAGAGGCTGCCAGTGAATTGCTC 259
Db 121 TCTTTTCACTGACCTTGTATCTTCTGCAATTAAGAGAGGCTGCCAGTGAATTGCTC 180
QY 260 TTTCTCGCTACATCTCGGATGGAAGCTTGAAGCTTGGCGCTGCTACCGGCTCACTGG 319
Db 181 TTTCTCGCTACATCTCGGATGGAAGCTTGAAGCTTGGCGCTGCTACCGGCTCACTGG 240
QY 320 TTTCACTCTCTGGAAGCTTGTATCTTCTGCAATTAAGAGAGGCTGCCAGTGAATTGCTC 379
Db 241 TTTCACTCTCTGGAAGCTTGTATCTTCTGCAATTAAGAGAGGCTGCCAGTGAATTGCTC 300
QY 380 AACCCCAACTCAGCTGAGAGATCTTACACCGCGGCTCTACTCTGTGAGAGACGCAAG 439
Db 301 AACCCCAACTCAGCTGAGAGATCTTACACCGCGGCTCTACTCTGTGAGAGACGCAAG 360
QY 440 GCTGAGCCGAGGGGCTGCGCGGCTGACCGCGCGGGGTGCAATATAGCCATCATGACC 499

```

Db 361 GCTGAGCCGAGGGGCTGCGCGGCTGACCGCGCGGGTGCAAAATGACATCATGACC 420  
QY 500 TTCAAGATTATTTTATCTGCTGGAATCTTTGTGAAAACATGAAAGACTTTCAA 559  
Db 421 TTCAAGATTATTTTATCTGCTGGAATCTTTGTGAAAACATGAAAGACTTTCAA 480  
QY 560 GCGTGGAGAGGGCTGATGAAATTCAGTTCGTCTCCAGACAGCTTGGCGCATCTT 619  
Db 481 GCGTGGAGAGGGCTGATGAAATTCAGTTCGTCTCCAGACAGCTTGGCGCATCTT 540  
QY 620 TTGCCCCGTATGAGTTGATGATTAAGACGACATTTGATCTTGGACATTTGA 676  
Db 541 TTGCCCCGTATGAGTTGATGATTAAGACGACATTTGATCTTGGACATTTGA 597

RESULT 12  
BT007402 597 bp mRNA linear PRI 13-MAY-2003  
LOCUS Homo sapiens activation-induced cytidine deaminase mRNA, complete cds.  
DEFINITION  
ACCESSION BT007402 GI:30583642  
VERSION  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 597)  
AUTHORS Kauline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Pheilan,M., and Farmer,A.  
TITLE Cloning of human full-length CDS in BD Creator(TM) System Donor vector

Unpublished  
2 (bases 1 to 597)  
Kauline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Pheilan,M., and Farmer,A.

Direct Submission  
Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA  
This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'Ac' after SalI site and before 'Atc' to provide Kozak consensus sequence; 'Gg' after last codon and before HindIII site to maintain reading frame.

Clone distribution: <http://bioinfo.clontech.com/orfclones>.  
Location/Qualifiers  
1..597  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="BD Creator(TM) CDS Library derived from MGC collection"  
/lab\_host="DH5alpha T1 resistant"  
/notes="vector: pDNR-Dual"  
1..597  
/codon\_start=1  
/product="activation-induced cytidine deaminase"  
/protein\_id="AAP36066.1"  
/db\_xref="GI:30583642"  
/translation="MDSLIMNRKFLYQFKNVMAKGRRETYLCVVRKDSANFSFLDRGLNKNKGCHVELFLFRYISDWLIDPGRCYRVMTFSWSPCYCAHVAADFLRGNPNSIRIATRIYFCEDRKAEPEGLRLHRAQVQIIMTFKDYFYCMNTEVENHERITFKAMEGLHENSVRISROLRILPLLYEVDLRDAFRLLG."

FEATURES  
source  
CDS  
ORIGIN  
Query Match 21.1%; Score 595; DB 9; Length 597;

Best Local Similarity 100.0%; Pred. No. 2, 6e-106;  
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 ATGACAGAGCTCTTGATGAACCGGAGAGTTTCTTACCAATTCAAAAATGTCGGTGG 139  
Db 1 ATGACAGAGCTCTTGATGAACCGGAGAGTTTCTTACCAATTCAAAAATGTCGGTGG 60  
QY 140 GCTAAGGTCGGCTGAGACCTTACCTGCTAGTATGAAAGAGGCTGACAGTCTACA 199  
Db 61 GCTAAGGTCGGCTGAGACCTTACCTGCTAGTATGAAAGAGGCTGACAGTCTACA 120  
QY 200 TCCTTTCACTGACCTTGGTTATCTTGCAATTAAGAACGGCTGCCAGTGAATTGGCTC 259  
Db 121 TCCTTTCACTGACCTTGGTTATCTTGCAATTAAGAACGGCTGCCAGTGAATTGGCTC 180  
QY 260 TTGCTCGGTACATCTGGAACCTGGAACCTGACCTGACCTGACCTGACCTGACCTG 319  
Db 181 TTGCTCGGTACATCTGGAACCTGGAACCTGACCTGACCTGACCTGACCTGACCTG 240  
QY 320 TTGACCTTCCTGAGACCTGCTGACCTGCTGACCTGCTGACCTGCTGACCTGCTG 379  
Db 241 TTGACCTTCCTGAGACCTGCTGACCTGCTGACCTGCTGACCTGCTGACCTGCTG 300  
QY 380 AACCCCAACCTGAGGATCTTACCGGCGGCTCTACTTGTGTGAGACCGCAAG 439  
Db 301 AACCCCAACCTGAGGATCTTACCGGCGGCTCTACTTGTGTGAGACCGCAAG 360  
QY 440 GCTGAGCCCGAGGGGCTGCGGCGCTGCAACCGGCGGCTGCAAAATGACATGATGACC 499  
Db 361 GCTGAGCCCGAGGGGCTGCGGCGCTGCAACCGGCGGCTGCAAAATGACATGATGACC 420  
QY 500 TTCAAGATTATTTTATCTGCTGGAATCTTTGTGAAAACATGAAAGACTTTCAA 559  
Db 421 TTCAAGATTATTTTATCTGCTGGAATCTTTGTGAAAACATGAAAGACTTTCAA 480  
QY 560 GCGTGGAGAGGGCTGATGAAATTCAGTTCGTCTCCAGACAGCTTGGCGCATCTT 619  
Db 481 GCGTGGAGAGGGCTGATGAAATTCAGTTCGTCTCCAGACAGCTTGGCGCATCTT 540  
QY 620 TTGCCCCGTATGAGTTGATGATTAAGACGACATTTGATCTTGGACATTTGA 674  
Db 541 TTGCCCCGTATGAGTTGATGATTAAGACGACATTTGATCTTGGACATTTGA 595

RESULT 13  
BT008226 597 bp mRNA linear SYN 13-MAY-2003  
LOCUS Synthetic construct Homo sapiens activation-induced cytidine deaminase mRNA, partial cds.  
DEFINITION  
ACCESSION BT008226 GI:30585290  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
TITLE  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Unpublished  
2 (bases 1 to 597)  
Kauline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., Labaer,J., Lin,Y., Pheilan,M., and Farmer,A.  
Direct Submission  
Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA  
This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal





```

Db      361 GCTGAGCCGAGGGGCTGCGGCGCTGCACCGCGCGGGGTGCAATAGCCATCATGACC 420
Qy      500 TTCAAGATTAATTTTCTCTCTGGAATCTTTGTAGAAAACCATGAAGAACTTTCAA 559
Db      421 TTCAAGATTAATTTTCTCTGGAATCTTTGTAGAAAACCATGAAGAACTTTCAA 480
Qy      560 GCTGAGGAGGGGCTGCATGAAAATTCAGTTCTCTCTCCAGACAGCTTGGGCATCTT 619
Db      481 GCTGAGGAGGGGCTGCATGAAAATTCAGTTCTCTCTCCAGACAGATTCGGGCATCTT 540
Qy      620 TTGCCCCGTATAGAGTTGATGACTTACGAGACGCAATTTGTACTTTGGGACTTTG 675
Db      541 TTGCCCCGTATAGAGTTGATGACTTACGAGACGCAATTTGTACTTTGGGACTTTG 596

RESULT 15
AF529816      596 bp      mRNA      linear      PRI 17-SEP-2002
LOCUS      AF529816
DEFINITION      Homo sapiens clone Ramos 2 AID (AID) mRNA, partial cds.
ACCESSION      AF529816
VERSION      AF529816.1 GI:22297219
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 596)
Martin, A. and Scharif, M.D.
Somatic hypermutation of the AID transgene in B and non-B cells
Proc. Natl. Acad. Sci. U.S.A. 99 (19), 12304-12308 (2002)
2 (bases 1 to 596)
Martin, A. and Scharif, M.D.
Direct Submision
Submitted (17-JUL-2002) Cell Biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave. Bronx, NY 10461, USA
FEATURES
source
1..596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="Ramos 2"
1..>596
/gene="AID"
1..>596
/feature="AID"
/note="Integrated into Burkitt's lymphoma cell line Ramos"
/codon_start=1
/product="AID"
/protein_id="PAM95403.1"
/db_xref="GI:22297220"
/translation="MDSILNRRKFLYQFANVWAKRRREYLYVVKRSDTSFSL
DFGLRNKNGCHGELFLFIRYISMDLDPGRCYRVFTSRSPCYDCARHVDLIRGND
NISLRITARYFCEDRKAPBGLRIHAGVOIAIMTKDYFCWNTFVENHETFK
AMGSLHENSVALSRQLRRLPLIEVDLIDAPRTLL"

ORIGIN
Query Match      21.1% Score 594.4; DB 9; Length 596;
Best Local Similarity 99.8%; Pred. No. 3.4e-106;
Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      80 ATGAGACGCTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAATGTCGCTG 139
Db      1 ATGAGACGCTCTTGATGAACCGAGAGAGTTCTTACCAATTCAAAATGTCGCTG 60
Qy      140 GCTAAGGCTGGCGCTGAGACCTTACCTGTCTAGTAGTAAGAGCGTGACAGTGTACA 199
Db      61 GCTAAGGCTGGCGCTGAGACCTTACCTGTCTAGTAGTAAGAGCGTGACAGTGTACA 120
Qy      200 TCCCTTTACCTGACCTTGGTTATCTTCGCAATPAGAAAGCGTCCAGCGTAATGCTC 259
Db      121 TCCCTTTACCTGACCTTGGTTATCTTCGCAATPAGAAAGCGTCCAGCGTAATGCTC 180
Qy      260 TTCTCCGCTACATCTCGAAGTGGAGCTGACCTGCGCTGTACCGCTCACCTTG 319

```

```

Db      181 TTCTCCGCTACATCTCGAAGTGGAGACTTACCCCTGCGCTCTACCGGTCACCTTG 240
Qy      320 TTCACTTCCTGAGACCCCTGCTACAGCTGTGCCCCGACATGTGAGCCGACTTTGCGAGGG 379
Db      241 TTCACTTCCTGAGACCCCTGCTACAGCTGTGCCCCGACATGTGAGCCGACTTTGCGAGGG 300
Qy      380 AACCCCACTCAGCTGAGAGATTTTCAACCGCGGCTCTACTCTGTAGAGACCGCAAG 439
Db      301 AACCCCACTCAGCTGAGAGATTTTCAACCGCGGCTCTACTCTGTAGAGACCGCAAG 360
Qy      440 GCTGAGCCGAGAGGGGCTGCGCGGCTGACACCGCGCGGGGTGCAATAGCATGATACC 499
Db      361 GCTGAGCCGAGAGGGGCTGCGCGGCTGACACCGCGCGGGGTGCAATAGCATGATACC 420
Qy      500 TTCAAGATTAATTTTCTCTGGAATCTTTGTAGAAAACCATGAAGAACTTTCAA 559
Db      421 TTCAAGATTAATTTTCTCTGGAATCTTTGTAGAAAACCATGAAGAACTTTCAA 480
Qy      560 GCTGAGGAGGGGCTGCATGAAAATTCAGTTCTCTCTCCAGACAGCTTGGGCATCTT 619
Db      481 GCTGAGGAGGGGCTGCATGAAAATTCAGTTCTCTCTCCAGACAGCTTGGGCATCTT 540
Qy      620 TTGCCCCGTATAGAGTTGATGACTTACGAGACGCAATTTGTACTTTGGGACTTTG 675
Db      541 TTGCCCCGTATAGAGTTGATGACTTACGAGACGCAATTTGTACTTTGGGACTTTG 596

```

Search completed: March 7, 2004, 09:57:47  
Job time : 10908 secs



PT Nucleic acid encoding activation induced cytidine deaminase, useful as a  
PT target for drug development for immune-related diseases including  
PT allergies.

PS Claim 3; Page 135-139; 174pp; Japanese.

XX  
XX The present sequence encodes human activation-induced cytidine deaminase  
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has  
CC cytidine activity similar to APOBEC-1. AID has antileukemic, antianemic,  
CC antitubercular, ophthalmological, anti-HIV and dermatological activities,  
CC and can be used in gene therapy. AID polynucleotides are useful in  
CC methods for identifying drugs for the treatment of B cell associated  
CC immune system disorders, immunodeficiency diseases and allergies, such as  
CC immunoglobulin A (Iga) deficiency disease, Iga nephritis, gamma-  
CC globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,  
CC drug allergy, allergic rhinitis, Rosen disease, DiGeorge disease, ataxia  
CC telangiectasia, common variable immunodeficiency disorder, MHC (major  
CC histocompatibility class II deficiency disease, AIDS (auto  
CC immunodeficiency syndrome), elevated IgE disorder, and IgG subclass  
CC selection disorder. The DNA sequences encoding AID may be used for gene  
CC therapy and the antibodies to the AID protein may be used for diagnosis  
CC and treatment of these disorders  
XX

Sequence 2818 BP; 868 A; 548 C; 626 G; 776 T; 0 U; 0 Other;

Query Match 100.0%; Score 2818; DB 3; Length 2818;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGACCATCATTAATGAGTGAATTTCTGGCCTGAGACTGAGGAGGAGAGA 60  
DB 1 AGAAGACCATCATTAATGAGTGAATTTCTGGCCTGAGACTGAGGAGGAGAGA 60  
QY 61 AGACACTCTGAGACCACTATGAGCAAGCCTCTTGATGAACCGAGAGAGATTCTTACCA 120  
DB 61 AGACACTCTGAGACCACTATGAGCAAGCCTCTTGATGAACCGAGAGAGATTCTTACCA 120  
QY 121 ATTCAAAATATGTCGCGGCTGAAGGTCGCGCTGAGACCTACTGTGCTAGTGTGA 180  
DB 121 ATTCAAAATATGTCGCGGCTGAAGGTCGCGCTGAGACCTACTGTGCTAGTGTGA 180  
QY 181 GAGGCGGTCAGTGTCACTATCTCTTCACTGGACTTGTATCTCGAATAAGAACG 240  
DB 181 GAGGCGGTCAGTGTCACTATCTCTTCACTGGACTTGTATCTCGAATAAGAACG 240  
QY 241 CTGCGACGTGGAATGCTCTCTCTCGCTACCTCTGAGACTGAGACCTGAGACCTGAGCG 300  
DB 241 CTGCGACGTGGAATGCTCTCTCTCGCTACCTCTGAGACTGAGACCTGAGACCTGAGCG 300  
QY 301 CTGCTACCGGTCACCTGCTTCACTCTGAGACCTCTGCTAGACTGCTGCGACATGT 360  
DB 301 CTGCTACCGGTCACCTGCTTCACTCTGAGACCTCTGCTAGACTGCTGCGACATGT 360  
QY 361 GGGCGGATTTCTGCGAGAGAGACCCCAACCTCACTGAGATCTTACCGCGGCTCTCA 420  
DB 361 GGGCGGATTTCTGCGAGAGAGACCCCAACCTCACTGAGATCTTACCGCGGCTCTCA 420  
QY 421 CTCTGTGAGGACCGCAGGCTGAGCCGAGGAGGCTGCGCGGCTGCAACGCGCGGCT 480  
DB 421 CTCTGTGAGGACCGCAGGCTGAGCCGAGGAGGCTGCGCGGCTGCAACGCGCGGCT 480  
QY 481 GCAATATGACATGATGATCTTCAAGATATTTTATCTGCTGAGATCTTTGTAGAAA 540  
DB 481 GCAATATGACATGATGATCTTCAAGATATTTTATCTGCTGAGATCTTTGTAGAAA 540  
QY 541 CCATGAAAGACCTTTCAAGACCTGAGAGGCTGCAATGAATTAAGTCTGCTCCAG 600  
DB 541 CCATGAAAGACCTTTCAAGACCTGAGAGGCTGCAATGAATTAAGTCTGCTCCAG 600  
QY 601 ACAGCTTGGGCGGATCTTTGGCCCTGATGAGGTGATGACTTACGAGAGCGCATTTG 660  
DB 601 ACAGCTTGGGCGGATCTTTGGCCCTGATGAGGTGATGACTTACGAGAGCGCATTTG 660

QY 661 TACTTGGGACCTTGTATGAGCACTTCCAGAGATGTCAACACGATGAAATATCTGCTG 720  
DB 661 TACTTGGGACCTTGTATGAGCACTTCCAGAGATGTCAACACGATGAAATATCTGCTG 720  
QY 721 AAGACAGTGAATAAAAACAGTCCCTCAAGTCTCTCTGTTTTATCTTCAACCTCTAC 780  
DB 721 AAGACAGTGAATAAAAACAGTCCCTCAAGTCTCTCTGTTTTATCTTCAACCTCTAC 780  
QY 781 TTTCTTGAAGTTTACAGAAAAATATTTATATACGACTCTTTAAAAAGATCTATGCTTG 840  
DB 781 TTTCTTGAAGTTTACAGAAAAATATTTATATACGACTCTTTAAAAAGATCTATGCTTG 840  
QY 841 AAAATAGAGAGAACACAGGTCGCGCAGAGAGCGCTCCATTTGCTGAGTTTGAAT 900  
DB 841 AAAATAGAGAGAACACAGGTCGCGCAGAGAGCGCTCCATTTGCTGAGTTTGAAT 900  
QY 901 GCAACATTTGCTCCCTACTGAGAAATACAGAACTGACAGACCTGAGACATCTTAAAGTGT 960  
DB 901 GCAACATTTGCTCCCTACTGAGAAATACAGAACTGACAGACCTGAGACATCTTAAAGTGT 960  
QY 961 CAACGTTTTCTATGACTTTTATGAGATGAGAGCAGAAAGTATCTTAAAAAGCATG 1020  
DB 961 CAACGTTTTCTATGACTTTTATGAGATGAGAGCAGAAAGTATCTTAAAAAGCATG 1020  
QY 1021 GTGAGAGATCAATGTTTTATATCAACATCTTTATTTATTTGATTCATTTGATTTAC 1080  
DB 1021 GTGAGAGATCAATGTTTTATATCAACATCTTTATTTATTTGATTCATTTGATTTAC 1080  
QY 1081 AGTGTGTATGATGATATTTTCTATTTCTTTCCCTGAGCTTTCCTTCAAGTATAC 1140  
DB 1081 AGTGTGTATGATGATATTTTCTATTTCTTTCCCTGAGCTTTCCTTCAAGTATAC 1140  
QY 1141 AAACCTTTCATCAAGGACATGATCTATAGACCTCTTATGAGAGATCTGAGTATGT 1200  
DB 1141 AAACCTTTCATCAAGGACATGATCTATAGACCTCTTATGAGAGATCTGAGTATGT 1200  
QY 1201 GACCCCAACCATCTCTCCCAAGATTAATCAATCAATGCGCTGATTTTAAATCAG 1260  
DB 1201 GACCCCAACCATCTCTCCCAAGATTAATCAATCAATGCGCTGATTTTAAATCAG 1260  
QY 1261 CAGAGAGATGTTTTATGTTTGTACAAAAAGATGTTATGAGTGGGAGTGAAGTATTA 1320  
DB 1261 CAGAGAGATGTTTTATGTTTGTACAAAAAGATGTTATGAGTGGGAGTGAAGTATTA 1320  
QY 1321 GACCATGATGATCACTTCAAGCTACTTAAATTAAGATCTTAAATGAGGACAGAGAC 1380  
DB 1321 GACCATGATGATCACTTCAAGCTACTTAAATTAAGATCTTAAATGAGGACAGAGAC 1380  
QY 1381 TGTGAACAGACACCTTAATATGAGTGTGATGTGATGAGCAATCTTCTGAAAACGC 1440  
DB 1381 TGTGAACAGACACCTTAATATGAGTGTGATGTGATGAGCAATCTTCTGAAAACGC 1440  
QY 1441 AAACCTTTTAAAGAACTCCCTAATTTAGAAAACCCCAACCTTCAATATCATATTA 1500  
DB 1441 AAACCTTTTAAAGAACTCCCTAATTTAGAAAACCCCAACCTTCAATATCATATTA 1500  
QY 1501 GCAAAACATTTGAAGAGTGTCTGGAATGTTGGGAGAGAGAAATCTATTTGCTCTGT 1560  
DB 1501 GCAAAACATTTGAAGAGTGTCTGGAATGTTGGGAGAGAGAAATCTATTTGCTCTGT 1560  
QY 1561 GGGTCTCTTCACTCAGAAATGCAAGTCAAGGTTGCTACATTTGTATGATGTGT 1620  
DB 1561 GGGTCTCTTCACTCAGAAATGCAAGTCAAGGTTGCTACATTTGTATGATGTGT 1620  
QY 1621 GATGCTTCTCCCAAGATATTTATCTATATAGAGATGTGACAAAACAGATATATTA 1680  
DB 1621 GATGCTTCTCCCAAGATATTTATCTATATAGAGATGTGACAAAACAGATATATTA 1680  
QY 1681 AGTGTGAAACGCTGACAGCTCATAGTCTGAGCTTGGAGAGGTTGAGAGAGAGGA 1740  
DB 1681 AGTGTGAAACGCTGACAGCTCATAGTCTGAGCTTGGAGAGGTTGAGAGAGAGGA 1740  
QY 1741 TGCGTTGAACAAGGTGTTCAAGGCGAGCTGAGCAACATTAACAAATCTGTCTCAA 1800

```
Db 1741 TGGCTTGAACACAGGTGTTCAAGGCCAGCCCTGGCCCAACATACAAAGATCCGTCTCTCA 1800
Qy 1801 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGGGGTGGTGGCTCAAGCCCTGTATCC 1860
Db 1801 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGGGGTGGTGGCTCAAGCCCTGTATCC 1860
Qy 1861 CAGCACTTTGGAGAGCCGAGCCGGGCGGATCACTGTGTCTCAAGAGTTTGAAGACAGCCT 1920
Db 1861 CAGCACTTTGGAGAGCCGAGCCGGGCGGATCACTGTGTCTCAAGAGTTTGAAGACAGCCT 1920
Qy 1921 GGGCCAACTGGCCAAACCCCGCTGTACTCAAAAATGCAAAATTAGCCAGCGGTGTAGC 1980
Db 1921 GGGCCAACTGGCCAAACCCCGCTGTACTCAAAAATGCAAAATTAGCCAGCGGTGTAGC 1980
Qy 1981 AGGCACTGTATATCCCACTACTTGGAGAGGTGAGGCGAGAGATCGCTTGAACCCAGGA 2040
Db 1981 AGGCACTGTATATCCCACTACTTGGAGAGGTGAGGCGAGAGATCGCTTGAACCCAGGA 2040
Qy 2041 GGTGAGAGGTGAGTAAGCTGAGATCGGCGCTTGCACTCCAGCCTGGGAGCAAGAGCA 2100
Db 2041 GGTGAGAGGTGAGTAAGCTGAGATCGGCGCTTGCACTCCAGCCTGGGAGCAAGAGCA 2100
Qy 2101 AGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db 2101 AGACTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Qy 2161 GGAAGAGAGAGATGGGGAGAGCATTGCAAGAAATGTCTTATCCAAAGATGTAGGA 2220
Db 2161 GGAAGAGAGAGATGGGGAGAGCATTGCAAGAAATGTCTTATCCAAAGATGTAGGA 2220
Qy 2221 GCCAATAGAGGATCCCATTTGCTCTTTGGTGTCTATTTGTCCTTAACTGTCTT 2280
Db 2221 GCCAATAGAGGATCCCATTTGCTCTTTGGTGTCTATTTGTCCTTAACTGTCTT 2280
Qy 2281 GACAGTGAAGAAATATTCAGAAATACCATATCCCTGTGCGGTATTAATTAAGCAACCT 2340
Db 2281 GACAGTGAAGAAATATTCAGAAATACCATATCCCTGTGCGGTATTAATTAAGCAACCT 2340
Qy 2341 TGCAATGAGATGAGAGATTCACAGAGAACTTGAATGCAACTGTCTTATTTATC 2400
Db 2341 TGCAATGAGATGAGAGATTCACAGAGAACTTGAATGCAACTGTCTTATTTATC 2400
Qy 2401 TTATTTGACATTAATTTGTAAAGATTAATTTGATTAATTTGATTAATTTGATTAAT 2460
Db 2401 TTATTTGACATTAATTTGTAAAGATTAATTTGATTAATTTGATTAATTTGATTAAT 2460
Qy 2461 TTATTTGATTTGGGTCTAATGATTTTATTAACATGATTTCTTTCTGATTAATTTGA 2520
Db 2461 TTATTTGATTTGGGTCTAATGATTTTATTAACATGATTTCTTTCTGATTAATTTGA 2520
Qy 2521 AATGAGTCTCAAGCTTCAATTAATTTAATTTAGAAATGATTTCAATTAACAGGTAT 2580
Db 2521 AATGAGTCTCAAGCTTCAATTAATTTAATTTAGAAATGATTTCAATTAACAGGTAT 2580
Qy 2581 GTAATTTGAACATTTGAGTAATGTTGTACGAAGCCATTTCTCTGATTTTGTAAACT 2640
Db 2581 GTAATTTGAACATTTGAGTAATGTTGTACGAAGCCATTTCTCTGATTTTGTAAACT 2640
Qy 2641 TTTATGACAGAAATTTGCTTGGCTCACTTTCAATCACTTAATTAATTAATTAATTAAT 2700
Db 2641 TTTATGACAGAAATTTGCTTGGCTCACTTTCAATCACTTAATTAATTAATTAATTAAT 2700
Qy 2701 TTTGAGAGCTGTGAAGATTAATTAACAAATTAATTAATTAATTAATTAATTAATTAAT 2760
Db 2701 TTTGAGAGCTGTGAAGATTAATTAACAAATTAATTAATTAATTAATTAATTAATTAAT 2760
Qy 2761 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2818
Db 2761 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2818
```

RESULT 2

```
AB573287
ID AB573287 standard; DNA; 2791 BP.
XX AC AB573287;
XX DT 04-DEC-2002 (first entry)
XX DE
XX DNA encoding human translocation del(12p) protein #2.
XX
XX Chromosome aberration; oncogenic fusion protein; cancer;
XX proliferative disease; cellular protein isoform; heat shock protein 90;
XX HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
XX T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CMV;
XX acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
XX acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
XX papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
XX rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
XX
XX Homo sapiens.
XX
XX WO200269900-A2.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002MO-US006518.
XX
XX 01-MAR-2001; 2001US-027251P.
XX
XX (CONF-) CONFORMA THERAPEUTICS CORP.
XX
XX Fritz LC, Burrows FU;
XX
XX WPI; 2002-698710/75.
XX
XX P-PSDB; AB595083.
XX
XX
XX Treating genetically-defined disease associated with chromosomal
XX aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
XX diseases, involves administering an inhibitor of heat shock protein 90.
XX
XX
XX Disclosure; Page 246-247; 389pp; English.
XX
XX The invention describes a method of treating genetically-defined disease
XX associated with chromosomal aberrations yielding oncogenic fusion
XX proteins (I), treating cancerous cells containing (II) in a heterogeneous
XX cell population, treating proliferative diseases associated with mutant
XX protein or cellular protein isoforms (II) dependent on heat shock protein
XX (HSP)-90, or selectively treating cells expressing (II) involving
XX administering HSP90-inhibitor. The method is useful for treating
XX genetically-defined disease with chromosomal aberration yielding
XX oncogenic fusion protein, treating cancerous cells containing fusion
XX protein in heterogeneous cell population, treating proliferative disease
XX (e.g. rheumatoid arthritis or cancer) associated with mutant protein or
XX cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.
XX p53), or selectively treating cells expressing mutant protein or cellular
XX protein isoform in a patient heterozygous for (II). The method is useful
XX for treating a disease e.g. haematopoietic disorder such as T or B cell
XX lymphoma, chronic myeloid leukaemia (CMV), APL, ALL, AML, NHL and CMML,
XX or a disease characterised by a solid tumour such as papillary thyroid
XX carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and
XX synovial sarcoma. The method is also useful for treating viral
XX infections. This represents the DNA sequence of a chromosome aberration
XX
XX Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;
XX
XX Query Match 99.0%; Score 2791; DB 6; Length 2791;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 2791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 GACCATGATTAATTAAGAGAGATTTTCTGCGCTGAGACTGACGAGGAGCAAGAGA 63
Db 1 GACCATGATTAATTAAGAGAGATTTTCTGCGCTGAGACTGACGAGGAGCAAGAGA 60
Qy 64 CACTCTGACACGACTATGACAGCCCTTGATGAAACCGAGAGAGTTCTTTACCAATT 123
```

Db 61 CACTGTGACACCACTATGACAGCCTCTGTATGAAACCGAGAACTTTCTTACCAAT 120  
Qy 124 CAAAATGTCGGCTAAGGGTCCGCGTGAAGACTTACTGTGTCTACTATGAAG 183  
Db 121 CAAAATGTCGGCTAAGGGTCCGCGTGAAGACTTACTGTGTCTACTATGAAG 180  
Qy 184 GCGTGAAGTGTCACTCTTTTCACTGAACTTTGTATCTTCCGAATPAGAACGGCTG 243  
Db 181 GCGTGAAGTGTCACTCTTTTCACTGAACTTTGTATCTTCCGAATPAGAACGGCTG 240  
Qy 244 CCACTGTAATGTCTTCTCTCCGCTAATCTCGAATGGAACCTTACCGCTG 303  
Db 241 CCACTGTAATGTCTTCTCTCCGCTAATCTCGAATGGAACCTTACCGCTG 300  
Qy 304 CTACCGGTGACCTGTGTCACTCTCTGAGCCCTGTGAGACTGTGCGGAGATGTGGC 363  
Db 301 CTACCGGTGACCTGTGTCACTCTCTGAGCCCTGTGAGACTGTGCGGAGATGTGGC 360  
Qy 364 CGACTTTCTGAGAGGAAACCCCAACCTCAGTCTGAGGATCTTCAACGCGGCTCTACT 423  
Db 361 CGACTTTCTGAGAGGAAACCCCAACCTCAGTCTGAGGATCTTCAACGCGGCTCTACT 420  
Qy 424 CTGTGAGAACCGCAAGGCTGAGCCGAGGGGCTGCGGCGGCTGACCGCGCGGCTGCA 483  
Db 421 CTGTGAGAACCGCAAGGCTGAGCCGAGGGGCTGCGGCGGCTGACCGCGCGGCTGCA 480  
Qy 484 AATGCAATCATGACCTTCAAGATTAATTTTAACTGTGGAATPACTTTGTAGAAACA 543  
Db 481 AATGCAATCATGACCTTCAAGATTAATTTTAACTGTGGAATPACTTTGTAGAAACA 540  
Qy 544 TGAAGAACTTTCAAAACCTGAGAGGGGCTGATGAAATTCAGTCTGTCTCCAGACA 603  
Db 541 TGAAGAACTTTCAAAACCTGAGAGGGGCTGATGAAATTCAGTCTGTCTCCAGACA 600  
Qy 604 GCTTCGGCGCATCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGAACGATTTGTAC 663  
Db 601 GCTTCGGCGCATCTTTTGGCCCTGTATGAGGTTGATGACTTACGAGAACGATTTGTAC 660  
Qy 664 TTTGGGACTTTGATGACACTTCCAGGAATGTCAACACGATGAATATCTGTCTGAAG 723  
Db 661 TTTGGGACTTTGATGACACTTCCAGGAATGTCAACACGATGAATATCTGTCTGAAG 720  
Qy 724 ACACTGATTAATAAAACGCTCTTCAAGTCTTCTGTCTTTTATTTCTCACTCTT 783  
Db 721 ACACTGATTAATAAAACGCTCTTCAAGTCTTCTGTCTTTTATTTCTCACTCTT 780  
Qy 784 CTTAGAGTTTACAGAAAAATTTATATAGACTCTTTAAAAAGATCTATGTCTGAAA 843  
Db 781 CTTAGAGTTTACAGAAAAATTTATATAGACTCTTTAAAAAGATCTATGTCTGAAA 840  
Qy 844 ATAGAGAGGAAACAGGCTGTGGCAGAGGAGCTGTGCAATGTGTGAGTTTGAATGA 903  
Db 841 ATAGAGAGGAAACAGGCTGTGGCAGAGGAGCTGTGCAATGTGTGAGTTTGAATGA 900  
Qy 904 ACATTTGCCCTTACTGGAATPACAGAACTGCAAGGACTGTGAGACTCTTAAGTGTCA 963  
Db 901 ACATTTGCCCTTACTGGAATPACAGAACTGCAAGGACTGTGAGACTCTTAAGTGTCA 960  
Qy 964 CGTTTCTTATGACTTTTATGATGAGAGCAAGGATGATGCTTAAGCAATGTG 1023  
Db 961 CGTTTCTTATGACTTTTATGATGAGAGCAAGGATGATGCTTAAGCAATGTG 1020  
Qy 1024 AGAGATCAATGTTTATATCAACATCTTATTTATTTGATTTGATTTGATTTGATTTG 1083  
Db 1021 AGAGATCAATGTTTATATCAACATCTTATTTATTTGATTTGATTTGATTTGATTTG 1080  
Qy 1084 GGTCTAGTATGATTTTCTATCTTTTCTTCCCTGAGTTTACTTCAATGATCAAA 1143  
Db 1081 GGTCTAGTATGATTTTCTATCTTTTCTTCCCTGAGTTTACTTCAATGATCAAA 1140  
Qy 1144 CTCTTCATGAGGCAATGATCTATAGACTCTTCAATGAGAGTATCTGGGTGATTTGTAC 1203

Db 1141 CTCTTCATGAGGCAATGATCTATAGACTCTTCAATGAGAGTATCTGGGTGATTTGTAC 1200  
Qy 1204 CCGAACAATCTCTCCAAAGATTAATTCGAATCAGCGCTGTATGTTTAACTACAG 1263  
Db 1201 CCGAACAATCTCTCCAAAGATTAATTCGAATCAGCGCTGTATGTTTAACTACAG 1260  
Qy 1264 AAGCATGTTTATGTTTGTACAAAAGAGATGTATGAGGTGGAGTGAAGTATGAC 1323  
Db 1261 AAGCATGTTTATGTTTGTACAAAAGAGATGTATGAGGTGGAGTGAAGTATGAC 1320  
Qy 1324 CATGATGTCACCTTCAAGTACTTAAATPAAAGATCTTAAATGGGACAGAGACTGT 1383  
Db 1321 CATGATGTCACCTTCAAGTACTTAAATPAAAGATCTTAAATGGGACAGAGACTGT 1380  
Qy 1384 GACCAAGCACCTCTAATATATGAGTGTATGTGAAAGTACCAATCTTCTGAAACGAAA 1443  
Db 1381 GACCAAGCACCTCTAATATATGAGTGTATGTGAAAGTACCAATCTTCTGAAACGAAA 1440  
Qy 1444 CTCTTTTAAAGAGTCCCTAATTTAGAAACCCCAACCTTCAATGATTAATGACA 1503  
Db 1441 CTCTTTTAAAGAGTCCCTAATTTAGAAACCCCAACCTTCAATGATTAATGACA 1500  
Qy 1504 AACATTTGAAAGAGTGTCTTGAATGTGGGAGAGAAATGTATGCTCTGAGG 1563  
Db 1501 AACATTTGAAAGAGTGTCTTGAATGTGGGAGAGAAATGTATGCTCTGAGG 1560  
Qy 1564 TCTCTCATCTCAGAAATGCAATCAGTCAAGGTTTGTCACTTTTGTATGTGTAT 1623  
Db 1561 TCTCTCATCTCAGAAATGCAATCAGTCAAGGTTTGTCACTTTTGTATGTGTAT 1620  
Qy 1624 GCTTCTCCAAAGATTAATTAATCTAATPAAAGTGTGACAAACAGAAATGATAAGC 1683  
Db 1621 GCTTCTCCAAAGATTAATTAATCTAATPAAAGTGTGACAAACAGAAATGATAAGC 1680  
Qy 1684 TGCAGAACGCTGACACGCTCATAGTCTAGCTCTTGGAGGTTGAGAGGAGATGG 1743  
Db 1681 TGCAGAACGCTGACACGCTCATAGTCTAGCTCTTGGAGGTTGAGAGGAGATGG 1740  
Qy 1744 CTTGAAACAGGTTTCAAGGCGCAGCTGTGGCAATPACAGATCTGTCTCAAAA 1803  
Db 1741 CTTGAAACAGGTTTCAAGGCGCAGCTGTGGCAATPACAGATCTGTCTCAAAA 1800  
Qy 1804 AAAAAAAAAAAAAAAAAAGAGAGAGAGGCGCGGCTGTGCTACGCTGTATTCAG 1863  
Db 1801 AAAAAAAAAAAAAAAAAAGAGAGAGAGGCGCGGCTGTGCTACGCTGTATTCAG 1860  
Qy 1864 CACTTTGGAGGCGGAGCGGCGGAGATCACTGTGTGTGAGAGTTTGAACCGAGCTGGC 1923  
Db 1861 CACTTTGGAGGCGGAGCGGCGGAGATCACTGTGTGTGAGAGTTTGAACCGAGCTGGC 1920  
Qy 1924 CAACATGCAAAACCCGCTGTATCTCAAAATGCAAAATTTAGCAGCGGTGTAGCAG 1983  
Db 1921 CAACATGCAAAACCCGCTGTATCTCAAAATGCAAAATTTAGCAGCGGTGTAGCAG 1980  
Qy 1984 CACTGTATCTCCAGCTATCTTGGAGGCTGAGGCAAGAAATGCTTGAACCCAGAGGT 2043  
Db 1981 CACTGTATCTCCAGCTATCTTGGAGGCTGAGGCAAGAAATGCTTGAACCCAGAGGT 2040  
Qy 2044 GAGGTTGCAATGAGCTGAGATGCTGTGCGTTTGCATCTCAGCTGGGGGAGCAAGCAAG 2103  
Db 2041 GAGGTTGCAATGAGCTGAGATGCTGTGCGTTTGCATCTCAGCTGGGGGAGCAAGCAAG 2100  
Qy 2104 CTCTGTCTCAGAAAAAAG 2163  
Db 2101 CTCTGTCTCAGAAAAAAG 2160  
Qy 2164 GAGAGAGTGGGAGAGATGCAAGAAATGTGCTTTATCCAAATGTAAAGAGCC 2223  
Db 2161 GAGAGAGTGGGAGAGATGCAAGAAATGTGCTTTATCCAAATGTAAAGAGCC 2220  
Qy 2224 AATAAGGATCCCTAATTTGTCTCTTTGTGTGTCTAATTTGTCTTAACTGTCTTGTAC 2283  
Db 2221 AATAAGGATCCCTAATTTGTCTCTTTGTGTGTCTAATTTGTCTTAACTGTCTTGTAC 2280

QY 2284 AGTGAAGAAAATATTGAGAAATACCATATCCCTGCGCTTATTACCTAGCAACCCCTTG 2343  
 Db 2281 AGTGAAGAAAATATTGAGAAATACCATATCCCTGCGCTTATTACCTAGCAACCCCTTG 2340  
 QY 2344 AATGAAGATGAGCAGATCCACAGGAAACTGGAATGACAACTGCTTATTATTATCTTA 2403  
 Db 2341 AATGAAGATGAGCAGATCCACAGGAAACTGGAATGACAACTGCTTATTATTATCTTA 2400  
 QY 2404 TTGACATTAAGTTGTAAGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 2463  
 Db 2401 TTGACATTAAGTTGTAAGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 2460  
 QY 2464 TATATTTTGGCTTAATGATTTTATTAACATGATTTCTTCTGATATATTTGAAT 2523  
 Db 2461 TATATTTTGGCTTAATGATTTTATTAACATGATTTCTTCTGATATATTTGAAT 2520  
 QY 2524 GGAGCTCAAGCTTCAATTAATTTATTAATTTAGAAATGATTTCTTAATTAACAAGTATGA 2583  
 Db 2521 GGAGCTCAAGCTTCAATTAATTTATTAATTTAGAAATGATTTCTTAATTAACAAGTATGA 2580  
 QY 2584 ATTGTAACATTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2643  
 Db 2581 ATTGTAACATTTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
 QY 2644 ATGACAGCAATTTGCTTCTGCTGCTCACTTCACTGATTAATTAATTAATTAATTTT 2703  
 Db 2641 ATGACAGCAATTTGCTTCTGCTGCTCACTTCACTGATTAATTAATTAATTAATTTT 2700  
 QY 2704 GGAAGCTGTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2763  
 Db 2701 GGAAGCTGTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760  
 QY 2764 AATTAATAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2794  
 Db 2761 AATTAATAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2791

RESULT 3  
 ABS73288  
 ID ABS73288 standard; DNA; 2791 BP.  
 XX  
 AC ABS73288;  
 XX  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE DNA encoding human translocation del(12p) protein #3.  
 XX  
 KM Chromosome aberration; oncogenic fusion protein; cancer;  
 KM Proliferative disease; cellular protein isoform; heat shock protein 90;  
 KM HSP-90; Rheumatoid arthritis; cancer; haematopoietic disorder;  
 KM T cell lymphoma; B cell lymphoma; chronic myeloid leukemia; CML;  
 KM acute myeloid leukemia; AML; chronic myelomonocytic leukemia; CMML;  
 KM papillary thyroid carcinoma; ALL; APL; NHL; solid tumour;  
 KM papillary thyroid carcinoma; Ewing's sarcoma; melanoma; rhabdomyosarcoma;  
 KM rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200269500-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 01-MAR-2002; 2002MO-US006518.  
 XX  
 PR 01-MAR-2001; 2001US-0272751P.  
 XX  
 PA (CONF-) CONFORMA THERAPEUTICS CORP.  
 XX  
 PI Filitz LC, Burows FU;  
 XX  
 DR WPI, 2002-688710/75.  
 DR P-PSDB; ABG95084.

XX Treating genetically-defined disease associated with chromosomal  
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative  
 diseases, involves administering an inhibitor of heat shock protein 90.  
 XX  
 XX Disclosure; Page 248-249; 389pp; English.  
 CC The invention describes a method of treating genetically-defined disease  
 CC associated with chromosomal aberrations yielding oncogenic fusion  
 CC proteins (I), treating cancerous cells containing (II) in a heterogeneous  
 CC cell population, treating proliferative diseases associated with mutant  
 CC protein or cellular protein isoforms (II) dependent on heat shock protein  
 CC (HSP)-90, or selectively treating cells expressing (II) involving  
 CC administering HSP90-inhibitor. The method is useful for treating  
 CC genetically-defined disease with chromosomal aberration yielding  
 CC oncogenic fusion protein, treating cancerous cells containing fusion  
 CC protein in heterogeneous cell population, treating proliferative disease  
 CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
 CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
 CC p53), or selectively treating cells expressing mutant protein or cellular  
 CC protein isoform in a patient heterozygous for (II). The method is useful  
 CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
 CC lymphoma, chronic myeloid leukemia (CML), APL, ALL, AML, NHL and CMML,  
 CC or a disease characterised by a solid tumour such as papillary thyroid  
 CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
 CC synovial sarcoma. The method is also useful for treating viral  
 CC infections. This represents the DNA sequence of a chromosome aberration  
 XX  
 SQ Sequence 2791 BP; 842 A; 548 C; 625 G; 776 T; 0 U; 0 Other;  
 Query Match 99.0%; Score 2791; DB 6; Length 2791;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 GAACATCATTAATTAAGTGAAGATTTTCTGCTGAGCTTGACAGGAGGCAAGAGA 63  
 Db 1 GAACATCATTAATTAAGTGAAGATTTTCTGCTGAGCTTGACAGGAGGCAAGAGA 60  
 QY 64 CACTCTGACACACATATGACAGCCTCTGATGATGATGATGATGATGATGATGATGATGAT 123  
 Db 61 CACTCTGACACACATATGACAGCCTCTGATGATGATGATGATGATGATGATGATGATGAT 120  
 QY 124 CAATAATGTCGCTGAGCTGAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 183  
 Db 121 CAATAATGTCGCTGAGCTGAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 180  
 QY 184 GCGTGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243  
 Db 181 GCGTGAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
 QY 244 CCAGTGAATTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303  
 Db 241 CCAGTGAATTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 304 CTACCGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
 Db 301 CTACCGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 QY 364 CGACTTTCTGAGAGGGAACCCCAACTCACTGATGATGATGATGATGATGATGATGATGAT 423  
 Db 361 CGACTTTCTGAGAGGGAACCCCAACTCACTGATGATGATGATGATGATGATGATGATGAT 420  
 QY 424 CTGTGAAGACCGGAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 483  
 Db 421 CTGTGAAGACCGGAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 480  
 QY 484 AATAGCATCATGATCTTCAAGATTTATTTATGCTGATGATGATGATGATGATGATGATGAT 543  
 Db 481 AATAGCATCATGATCTTCAAGATTTATTTATGCTGATGATGATGATGATGATGATGATGAT 540  
 QY 544 TGAAGAACTTTCAAGGCTGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGAT 603  
 Db 541 TGAAGAACTTTCAAGGCTGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGAT 600

QY 604 GCTTCGGCGCATCTTTTGGCCCTGTATGAGTGTGATGACTTACGAGACGATTCGTGAC 663  
DB 601 GCTTCGGCGCATCTTTTGGCCCTGTATGAGTGTGATGACTTACGAGACGATTCGTGAC 660  
QY 664 TTTGGGACTTTGATGACCACTTCAGGAATGTGACACAGATGAATCTCTGTGAAG 723  
DB 661 TTTGGGACTTTGATGACCACTTCAGGAATGTGACACAGATGAATCTCTGTGAAG 720  
QY 724 ACAGTGTATAAAAACAGTCTTCAAGTCTTCTGTGTTTATCTTCACTCTCACTT 783  
DB 721 ACAGTGTATAAAAACAGTCTTCAAGTCTTCTGTGTTTATCTTCACTCTCACTT 780  
QY 784 CTAGAGTTTACAGAAAAATTTTATATATGACTCTTTAAAAAGATCTATGCTTGAA 843  
DB 781 CTAGAGTTTACAGAAAAATTTTATATGACTCTTTAAAAAGATCTATGCTTGAA 840  
QY 844 ATAGAGAGAGACAGAGTCTGCGCAGAGACGTGTGCAATTTGTCAGATTTGATGCA 903  
DB 841 ATAGAGAGAGACAGAGTCTGCGCAGAGACGTGTGCAATTTGTCAGATTTGATGCA 900  
QY 904 ACATGTCCCTTCTGAGGAATTAACAGACTGACAGACTTGGAGACATCTTAAAGTGTCA 963  
DB 901 ACATGTCCCTTCTGAGGAATTAACAGACTGACAGACTTGGAGACATCTTAAAGTGTCA 960  
QY 964 CGTTTTCTATGACTTTTAGTAGATGAGAGACAGAGATGATCTTAAAAAGCATGTG 1023  
DB 961 CGTTTTCTATGACTTTTAGTAGATGAGAGACAGAGATGATCTTAAAAAGCATGTG 1020  
QY 1024 AGAGATCAATGTTTTTATATCAACATCTTATATTTGATTTGATTTGATTTGATTTA 1083  
DB 1021 AGAGATCAATGTTTTTATATCAACATCTTATATTTGATTTGATTTGATTTGATTTA 1080  
QY 1084 GGTGTATGATGATTTTCTATCTTTTCCCTTACGTTTACTTCAATGACACAA 1143  
DB 1081 GGTGTATGATGATTTTCTATCTTTTCCCTTACGTTTACTTCAATGACACAA 1140  
QY 1144 CTCTTCATCAAGCCATGATCTATAGGACTCTCTAATGAGATCTTGGGTGATGTGAC 1203  
DB 1141 CTCTTCATCAAGCCATGATCTATAGGACTCTCTAATGAGATCTTGGGTGATGTGAC 1200  
QY 1204 CCCAAACCATCTGCCAAAGCATTAATATCCATCATGCGGTATGTTTATACAGAC 1263  
DB 1201 CCCAAACCATCTGCCAAAGCATTAATATCCATCATGCGGTATGTTTATACAGAC 1260  
QY 1264 AAGCATGTTTTTATGTTTGTACAAAAGATGTTATGGGTGGGATGAGGTATAGAC 1323  
DB 1261 AAGCATGTTTTTATGTTTGTACAAAAGATGTTATGGGTGGGATGAGGTATAGAC 1320  
QY 1324 CATGATGTGACCTTCAAGCTTATTAATTAAGATCTTAAATGGGACAGAGACTGT 1383  
DB 1321 CATGATGTGACCTTCAAGCTTATTAATTAAGATCTTAAATGGGACAGAGACTGT 1380  
QY 1384 GATCAACACACCCCTAATTAATGAGGTGATGTCTGAATAGCAAACTTCTGGAAGCGCAA 1443  
DB 1381 GATCAACACACCCCTAATTAATGAGGTGATGTCTGAATAGCAAACTTCTGGAAGCGCAA 1440  
QY 1444 CTCTTTAAGGAAGTCCCTAATTTAGAAACACCAACCACTTCAATCATATATTAGA 1503  
DB 1441 CTCTTTAAGGAAGTCCCTAATTTAGAAACACCAACCACTTCAATCATATATTAGA 1500  
QY 1504 AACAAITGGAAGAGTGTGTTGATGATTTGGGAGAGAAATCTATTTGCTCTGTGG 1563  
DB 1501 AACAAITGGAAGAGTGTGTTGATGATTTGGGAGAGAAATCTATTTGCTCTGTGG 1560  
QY 1564 TCTCTTATCTCGAAATGCGCATCGGTCTAGGTTTCTCACTTTGTATGTGTAT 1623  
DB 1561 TCTCTTATCTCGAAATGCGCATCGGTCTAGGTTTCTCACTTTGTATGTGTAT 1620  
QY 1624 GCTTCTCCCAAGGTATTTACTATATTAAGAGTTGTGACAAAACAGATGATTAAGC 1683  
DB 1621 GCTTCTCCCAAGGTATTTACTATATTAAGAGTTGTGACAAAACAGATGATTAAGC 1680

QY 1684 TGCACCGTGGCACAGCTCATAGTTCTAGCTGTGGAGGTTGAGAGAGATGG 1743  
DB 1681 TGCACCGTGGCACAGCTCATAGTTCTAGCTGTGGAGGTTGAGAGAGATGG 1740  
QY 1744 CTGACACAGTGTTCAGGCCAGCTGGGCAACATTAACAATCTGTCTCTCAAAA 1803  
DB 1741 CTGACACAGTGTTCAGGCCAGCTGGGCAACATTAACAATCTGTCTCTCAAAA 1800  
QY 1804 AAAAAAAAAAAAAAGAGAGAGGCGGGGCTGTGTGCTCAGGCTGTAACTCCAG 1863  
DB 1801 AAAAAAAAAAAAAAGAGAGAGGCGGGGCTGTGTGCTCAGGCTGTAACTCCAG 1860  
QY 1864 CACTTGGAGGCGGAGCGGAGATCACTGTGTGCTCAGAGTTTGAACAGCCTGGC 1923  
DB 1861 CACTTGGAGGCGGAGCGGAGATCACTGTGTGCTCAGAGTTTGAACAGCCTGGC 1920  
QY 1924 CAACATGGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAAGCGGTGTAGAG 1983  
DB 1921 CAACATGGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAAGCGGTGTAGAG 1980  
QY 1984 CACTGTATCTCCAGCTTCTTTGGAGGCTGAGGACAGAGATGCTTGAACCAAGAGT 2043  
DB 1981 CACTGTATCTCCAGCTTCTTTGGAGGCTGAGGACAGAGATGCTTGAACCAAGAGT 2040  
QY 2044 GGAGTTGACGTAGCTGATGATGTGCGCTTGAATCCAGCCTGGGCGACAGAGCAAG 2103  
DB 2041 GGAGTTGACGTAGCTGATGATGTGCGCTTGAATCCAGCCTGGGCGACAGAGCAAG 2100  
QY 2104 CTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2163  
DB 2101 CTCTGTCTCAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
QY 2164 GAGAGAGTGGGAGAGATGCAAGAGAAATGTGCTTTATCCAAACAAATGTAGAGCC 2223  
DB 2161 GAGAGAGTGGGAGAGATGCAAGAGAAATGTGCTTTATCCAAACAAATGTAGAGCC 2220  
QY 2224 AATAGGATGCCATTTGTCTCTTTGGTGTCTATTTGCCCTAACAATGCTTTGAC 2283  
DB 2221 AATAGGATGCCATTTGTCTCTTTGGTGTCTATTTGCCCTAACAATGCTTTGAC 2280  
QY 2284 AGTGAAGAAATATTCAGAAATTAACATATCCCTGTGCGTTATTAACCAACCTTGC 2343  
DB 2281 AGTGAAGAAATATTCAGAAATTAACATATCCCTGTGCGTTATTAACCAACCTTGC 2340  
QY 2344 AATGAAGTGAAGATCCACAGAGAAACTTGAATGACAACTGTCTTATTTAATCTTA 2403  
DB 2341 AATGAAGTGAAGATCCACAGAGAAACTTGAATGACAACTGTCTTATTTAATCTTA 2400  
QY 2404 TTGTACATTAAGTTGTAAAGATTTAAATTTGTACTGTATGATTTATATTTTA 2463  
DB 2401 TTGTACATTAAGTTGTAAAGATTTAAATTTGTACTGTATGATTTATATTTTA 2460  
QY 2464 TATATTTTGGTATATGATTTTATTAATGAATTTCTTTCGATATATGAAT 2523  
DB 2461 TATATTTTGGTATATGATTTTATTAATGAATTTCTTTCGATATATGAAT 2520  
QY 2524 GAGTCTCAAGCTTCAATTAATTTATTAATTTAGAAATGATCTTAATCAACGATGTA 2583  
DB 2521 GAGTCTCAAGCTTCAATTAATTTATTAATTTAGAAATGATCTTAATCAACGATGTA 2580  
QY 2584 ATTGTAACTTGAAGTATAGTGTGACAAAGCAATTTCTTGAATTTTATTAATCTTT 2643  
DB 2581 ATTGTAACTTGAAGTATAGTGTGACAAAGCAATTTCTTGAATTTTATTAATCTTT 2640  
QY 2644 ATGACAGCAATTTGCTTGTGCTCACTTCAATCGTTAAATTAATGAATTAATTTT 2703  
DB 2641 ATGACAGCAATTTGCTTGTGCTCACTTCAATCGTTAAATTAATGAATTAATTTT 2700  
QY 2704 GGAAGCTGTGAAGTAAATACCAAAATTAATTAATTAATTAATTAATTAATTAAT 2763  
DB 2701 GGAAGCTGTGAAGTAAATACCAAAATTAATTAATTAATTAATTAATTAATTAAT 2760  
QY 2764 AATTAATAATCAATATGATGAATTAATCTTG 2794



Db 2761 AATATAATCAGATATGATGATTAACCTTG 2791

## RESULT 4

ID AAC5314 standard; DNA; 6564 BP.

XX AAC5314;

XX 05-FEB-2001 (first entry)

XX Human activation-induced cytidine deaminase genomic DNA SEQ ID NO:10.

XX Activation-induced cytidine deaminase; AID; cytidine deaminase;  
XX immune related disease; allergy; allergic disease; anti-allergic;  
XX antidiabetic; ophthalmologic; anti-HIV; dermatologic;  
XX gene therapy; B cell associated immune system disorder; food allergy;  
XX immunodeficiency disease; immunoglobulin A deficiency disease; asthma;  
XX IGA nephritis; gamma-globulinemia; atopic dermatitis; allergic colitis;  
XX drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;  
XX ataxia telangiectasia; common variable immunodeficiency disorder;  
XX major histocompatibility class II deficiency disease;  
XX auto immunodeficiency syndrome; IgG subclass selection disorder; ds.

XX Homo sapiens.

XX WO200058480-A1.

XX 05-OCT-2000.

XX 28-MAR-2000; 2000WO-JP001918.

XX 29-MAR-1999; 99JP-00087192.

XX 24-JUN-1999; 99JP-00178999.

XX 27-DEC-1999; 99JP-00371382.

XX (NIBS) JAPAN TOBACCO INC.

XX (HONO/) HONO T.

XX Honjo T, Muramatsu M;

XX WPI; 2000-611715/58.

XX Claim 17; Page 145-150; 174pp; Japanese.

XX The present invention describes an activation-induced cytidine deaminase  
XX (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has  
XX cytidine activity similar to APOBEC-1. AID has anti-allergic, anti-anemic,  
XX antidiabetic, ophthalmologic, anti-HIV and dermatological activities,  
XX and can be used in gene therapy. AID polynucleotides are useful in  
XX methods for identifying drugs for the treatment of B cell associated  
XX immune system disorders, immunodeficiency diseases and allergies, such as  
XX immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-  
XX globulinemia, atopic dermatitis, allergic colitis, asthma, food allergy,  
XX drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia  
XX telangiectasia, common variable immunodeficiency disorder, MHC (major  
XX histocompatibility class) class II deficiency disease, AIDS (auto  
XX immunodeficiency syndrome), elevated IGE disorder, and IGE subclass  
XX selection disorder. The DNA sequences encoding AID may be used for gene  
XX therapy and the antibodies to the AID protein may be used for diagnosis  
XX and treatment of these disorders. The present sequence represents a  
XX genomic DNA sequence of human AID

XX Sequence 6564 BP; 1909 A; 1358 C; 1383 G; 1914 T; 0 U; 0 Other;

XX Query Match 77.2%; Score 2174.6; DB 3; Length 6564;

XX Best Local Similarity 99.8%; Pired. No. 0;

XX Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 617 CTTTGGCCCTGTATGAGGTTGATGACTTACGAGACGATTTGCTTGGACTTTGA 676  
DB 3735 CTACAGCCCTGTATGAGGTTGATGACTTACGAGACGATTTGCTTGGACTTTGA 3794  
QY 677 TAGCACTTCCAGAAATGTCACACAGATGAAATATCTGCTGGAAGACAGTGTATAAA 736  
DB 3795 TAGCACTTCCAGAAATGTCACACAGATGAAATATCTGCTGGAAGACAGTGTATAAA 3854  
QY 737 AACAGTCTTCAAGCTCTCTCTGTTTATTTTCACTCACTCTTCTTGAAGTTTACA 796  
DB 3855 AACAGTCTTCAAGCTCTCTCTGTTTATTTTCACTCACTCTTCTTGAAGTTTACA 3914  
QY 797 GAAAAATATTTATATACGACTCTTAAAAAGATCTATGCTTGAANAATAGAGAAGAC 856  
DB 3915 GAAAAATATTTATATACGACTCTTAAAAAGATCTATGCTTGAANAATAGAGAAGAC 3974  
QY 857 AACAGTCTGCGCAGAGGACGCTGCAATGCTGAGTTTGAATGCAACATGTCCTTA 916  
DB 3975 AACAGTCTGCGCAGAGGACGCTGCAATGCTGAGTTTGAATGCAACATGTCCTTA 4034  
QY 917 CTGGGAATTAACAGAACTGACAGAACTTGGAGACATCTTAAAGTGCATTTTCTATGA 976  
DB 4035 CTGGGAATTAACAGAACTGACAGAACTTGGAGACATCTTAAAGTGCATTTTCTATGA 4094  
QY 977 CTTTAAAGTATGAGAGCAGAAAGTATGCTTAAAAAGCATGCTGAGATCAATG 1036  
DB 4095 CTTTAAAGTATGAGAGCAGAAAGTATGCTTAAAAAGCATGCTGAGATCAATG 4154  
QY 1037 TTTTATATCAACATCTTTATTTTGAATTCATTGATTAACAGTGTGTTATGATATA 1096  
DB 4155 TTTTATATCAACATCTTTATTTTGAATTCATTGATTAACAGTGTGTTATGATATA 4214  
QY 1097 GATTTTCTATCTTTTCCCTGAGCTTCTTCAAGTAAACAACCTCTTCAATGAG 1156  
DB 4215 GATTTTCTATCTTTTCCCTGAGCTTCTTCAAGTAAACAACCTCTTCAATGAG 4274  
QY 1157 CCATGATCTATAGAGACCTCTTAATGAGATATCTGGTGAATGGAACCCAAACATCTC 1216  
DB 4275 CCATGATCTATAGAGACCTCTTAATGAGATATCTGGTGAATGGAACCCAAACATCTC 4334  
QY 1217 TCCAAAGCATTAATATCCATCATGCTGATGTTTAAATCAGCAGAGCATGTTTAA 1276  
DB 4335 TCCAAAGCATTAATATCCATCATGCTGATGTTTAAATCAGCAGAGCATGTTTAA 4394  
QY 1277 TGTTTGTAACAAAAGATGTTATGAGTGGAGATGAGATTAACCATGATGTCAC 1336  
DB 4395 TGTTTGTAACAAAAGATGTTATGAGTGGAGATGAGATTAACCATGATGTCAC 4454  
QY 1337 CTCAAGCTACTTAAATAAAGATCTTAAATGGGACAGAGCTGGAACAAGACACC 1396  
DB 4455 CTCAAGCTACTTAAATAAAGATCTTAAATGGGACAGAGCTGGAACAAGACACC 4514  
QY 1397 TAATATAGGTTGATGCTGTAAGTACAAATCTTCTGAAAACGAAACTTTTAAAGAA 1456  
DB 4515 TAATATAGGTTGATGCTGTAAGTACAAATCTTCTGAAAACGAAACTTTTAAAGAA 4574  
QY 1457 GTCCCTAATTTGAAAACCCCAAACTTCAATATATTAATGCAACAAATTGGAAG 1516  
DB 4575 GTCCCTAATTTGAAAACCCCAAACTTCAATATATTAATGCAACAAATTGGAAG 4634  
QY 1517 AAGTGTGTAATGTTGGGAGAGAAATCTATGCTGCTGAGGCTCTTCAATCTCA 1576  
DB 4635 AAGTGTGTAATGTTGGGAGAGAAATCTATGCTGCTGAGGCTCTTCAATCTCA 4694  
QY 1577 GAAATGCCAATCAGATCAGATTTGCTACATTTTGTATGTTGATGCTTCCCAAG 1636  
DB 4695 GAAATGCCAATCAGATCAGATTTGCTACATTTTGTATGTTGATGCTTCCCAAG 4754  
QY 1637 GTATATTAATCTATATTAAGAGATTTGTAACAAAACAGATATTAAGCTGGAACCGTGGC 1696  
DB 4755 GTATATTAATCTATATTAAGAGATTTGTAACAAAACAGATATTAAGCTGGAACCGTGGC 4814

QY	1697	CAAGCCTATAGTCTTACTGCTCTGGAGAGTTGAGGAGAGAGATGGCTTGAACAAGT	1756
Db	4815	ACAAGCTATAGTCTTACTGCTCTGGAGAGTTGAGGAGAGATGGCTTGAACAAGT	4874
QY	1757	GTTCAAGGCCAGCCTGGGCAACATACAGAATCCGTGCTCTCAAAAAAAAAAAAAAAAAA	1816
Db	4875	GTTCAAGGCCAGCCTGGGCAACATACAGAATCCGTGCTCTCAAAAAAAAAAAAAAAAAA	4934
QY	1817	AAGAAAGAGAGAGGCGCGGCGTGCTCAAGCTGTAAATCCAGACCTTTGGAGGC	1876
Db	4935	AAGAAAGAGAGAGGCGCGGCGTGCTCAAGCTGTAAATCCAGACCTTTGGAGGC	4994
QY	1877	CGAGCGGGGGAGATCACTGTGGTCAAGAGTTTGAGACCAAGCTGGGCAATGGCAAA	1936
Db	4995	CGAGCGGGGGAGATCACTGTGGTCAAGAGTTTGAGACCAAGCTGGGCAATGGCAAA	5054
QY	1937	CCCGCTGTACTCAAAATGCAAAATATGACAGCGCGTATGAGAGCACTGTATCC	1996
Db	5055	CCCGCTGTACTCAAAATGCAAAATATGACAGCGCGTATGAGAGCACTGTATCC	5114
QY	1997	AGCTACTTGGAGGCTGAGGCAAGAAATCGCTTAAATCCAGAGGTGAGGTTGACATA	2056
Db	5115	AGCTACTTGGAGGCTGAGGCAAGAAATCGCTTAAATCCAGAGGTGAGGTTGACATA	5174
QY	2057	AGCTGAGATCGTCCGCTTGCACTCCAGCCTGGGCGACAAAGCAAGATCTGTCTCAGAA	2116
Db	5175	AGCTGAGATCGTCCGCTTGCACTCCAGCCTGGGCGACAAAGCAAGATCTGTCTCAGAA	5234
QY	2117	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAAAGAACATATTTGGGAGAGAAAGATGGG	2176
Db	5235	AAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAAAGAACATATTTGGGAGAGAGATGGG	5294
QY	2177	AAGCATTGCAAGAAATTTGCTTTATCCAAACAAATGTAAAGACCAATPAGGATCCC	2236
Db	5295	AAGCATTGCAAGAAATTTGCTTTATCCAAACAAATGTAAAGACCAATPAGGATCCC	5354
QY	2237	TATTTGTCTCTTTGGTGTCTATTGTGCCCTAACACATGCTCTTGAACGTAGAAAAATA	2296
Db	5355	TATTTGTCTCTTTGGTGTCTATTGTGCCCTAACACATGCTCTTGAACGTAGAAAAATA	5414
QY	2297	TTGAGAAATACATATCCCTGGCGGTATTAACGACAGCCCTGCAATGAAGTAGGC	2356
Db	5415	TTGAGAAATACATATCCCTGGCGGTATTAACGACAGCCCTGCAATGAAGTAGGC	5474
QY	2357	AGATCCACAGAAAACTTGAATGCAACAATGCTTATTTTATCTTATTTGTACATAGTT	2416
Db	5475	AGATCCACAGAAAACTTGAATGCAACAATGCTTATTTTATCTTATTTGTACATAGTT	5534
QY	2417	TGTAAAGAGTTAAAAATTTGTTACTCTCAATGATATCATTTATATTTATTTTGGCT	2476
Db	5535	TGTAAAGAGTTAAAAATTTGTTACTCTCAATGATATCATTTATATTTATTTTGGCT	5594
QY	2477	CTAATGATTTTTTATTAACGATTTCCCTTTTCGATATATTTGAAATGCAAGC	2536
Db	5595	CTAATGATTTTTTATTAACGATTTCCCTTTTCGATATATTTGAAATGCAAGC	5654
QY	2537	TTCATTAATTTATTAATCTTAGAAATGATTTCAATACACAGATATGTAATTTGAACATTC	2596
Db	5655	TTCATTAATTTATTAATCTTAGAAATGATTTCAATACACAGATATGTAATTTGAACATTC	5714
QY	2597	AGTAATGGTCTAGCAGAGCCATTTCTGTGATTTTATAGTAACTTTATGACAGAAAT	2656
Db	5715	AGTAATGGTCTAGCAGAGCCATTTCTGTGATTTTATAGTAACTTTTATGACAGAAAT	5774
QY	2657	TGCTCTGGGCTCACTTGAATCAGTTAATGTAATGATTAATTAATTTTGGAGCTGTGAG	2716
Db	5775	TGCTCTGGGCTCACTTGAATCAGTTAATGTAATGATTAATTAATTTTGGAGCTGTGAG	5834
QY	2717	ATTAATAATCAAAATATAATATATAAAGTGATTTATATGAAGTTAAATATAAATCAG	2776
Db	5835	ATTAATAATCAAAATATAATATATAAAGTGATTTATATGAAGTTAAATATAAATCAG	5894
QY	2777	TATGATGAAATPAACTTGAAA	2797

[illegible]

QY	617	CTTTTGCCCTCTGATGAGGTGATGACCTTAGAGACGATTTGCTACTTTGGGACTTTGA	676
Db	8950	CTACAGCCCCCTGTATGAGGTGATGATCTTAGAGACGATTTGCTACTTTGGGACTTTGA	9009
QY	677	TAGCACTCTCCAGGAATGTCACACAGATGAAATCTCTGCTGAAAGCAGTGGATGAAA	736
Db	9010	TAGCAACTCTCCAGGAATGTCACACAGATGAAATCTCTCTGTAAGCAGTGGATGAAA	9069
QY	737	AAACGTCCTCAAGTCTTCTGTTTTATTCCTCACTCTCACTTTCTTAGATTACA	796
Db	9070	AACAGTCCTCAAGTCTTCTGTTTTATTCCTCACTCTCACTTTCTTAGATTACA	9129
QY	797	GAATAAATATTTATATAGACTCTTTAAAAAGATCTATGCTTGAATAAGAGGAAC	856
Db	9130	GAATAAATATTTATATAGACTCTTTAAAAAGATCTATGCTTGAATAAGAGGAAC	9189
QY	857	ACAGGTCGGCCAGGAGAGTGCTGCAATGGAGAGTTTGAATGCAACATTGGCCCTTA	916
Db	9190	ACAGGTCGGCCAGGAGAGTGCTGCAATGGAGAGTTTGAATGCAACATTGGCCCTTA	9249
QY	917	CTGGAAATACAGAACTGCAGGACCTGGGAGCATCTTAAAGTGTCAACGTTTTCTATGA	976
Db	9250	CTGGAAATACAGAACTGCAGGACCTGGGAGCATCTTAAAGTGTCAACGTTTTCTATGA	9309
QY	977	CTTTTAGTAGGATGAGACAGAAAGTATGATCTTAAAAAGCATGGTAGAGATCAATG	1036
Db	9310	CTTTTAGTAGGATGAGACAGAAAGTATGATCTTAAAAAGCATGGTAGAGATCAATG	9369
QY	1037	TTTTATATTCAAACCTCTTATATTTATGATTCATTTGAGTTAAAGGTGTTAGTATA	1096
Db	9370	TTTTATATTCAAACCTCTTATATTTATGATTCATTTGAGTTAAAGGTGTTAGTATA	9429
QY	1097	GATTTTCTATCTTTTCCCTTGACGTTTACTTCAAGTAAACAACAACCTGTCATGAG	1156
Db	9430	GATTTTCTATCTTTTCCCTTGACGTTTACTTCAAGTAAACAACAACCTGTCATGAG	9489
QY	1157	CCATGATCTATAGGACCTCTTAATGAGATCTGGGTGATTTGACCCCAACCATCTC	1216
Db	9490	CCATGATCTATAGGACCTCTTAATGAGATCTGGGTGATTTGACCCCAACCATCTC	9549
QY	1217	TCGAAAGATTAATATCCATCATAGCGCTGATGTTTTAATCAGCAGAGCATGTTTTTA	1276
Db	9550	TCGAAAGATTAATATCCATCATAGCGCTGATGTTTTAATCAGCAGAGCATGTTTTTA	9609
QY	1277	TGTTTGTACAAAGAAAGTGTATAGGTGGGGAGTAGAGATATGACATGCTATGAC	1336
Db	9610	TGTTTGTACAAAGAAAGTGTATAGGTGGGGAGTAGAGATATGACATGCTATGCTAC	9669
QY	1337	CTTCAAGCTACTTTAATTAAGATCTTAAAAAGGCGACAGAGCATGTGAACAACACC	1396
Db	9670	CTTCAAGCTACTTTAATTAAGATCTTAAAAAGGCGACAGAGCATGTGAACAACAACC	9729
QY	1337	TAAATATGGGTGATGTCGTGAAGTAGCAATCTTCTGAAAACGCAACTCTTTTAAAGAA	1456
Db	9730	TAAATATGGGTGATGTCGTGAAGTAGCAAACTCTCTGAAAACGCAACTCTTTTAAAGAA	9789
QY	1457	GTCCTTAATTTAGAAACACCCACAACCTTCAATATATATTTAGCAAAACAATTGGAAG	1516
Db	9790	GTCCTTAATTTAGAAACACCCACAACCTTCAATATATATTTAGCAAAACAATTGGAAG	9849
QY	1517	AAGTTGCTTGATGTTGGGAGAGAAATCTATGAGCTCTCGTGGGTCTCTCATCTCA	1576
Db	9850	AAGTTGCTTGATGTTGGGAGAGAAATCTATGAGCTCTCGTGGGTCTCTCATCTCA	9909
QY	1577	GAAAATGCCAATCAGGTCAGAGTTGTCTACATTTTGTATGTGTGTGATGCTTCCCAAG	1636
Db	9910	GAAAATGCCAATCAGGTCAGAGTTGTCTACATTTTGTATGTGTGTGATGCTTCCCAAG	9969
QY	1637	GTAATTTAATCATAATTAAGAGTTGTGACAAAACAAGATATTAAGTGTGGCAACGCTGGC	1696
Db	9970	GTAATTTAATCATAATTAAGAGTTGTGACAAAACAAGATATTAAGTGTGGCAACGCTGGC	10029

[illegible]



```

Db      10030 ACAGCTCATAGTTAGCTGTGGAGGTTGAGGAGGAGATGGCTTGAACAGGCT 10089
Qy      1757 GTTCAAGGCCAGCTGGGCACTAACAGATCTCTCTCAAAAAAAAAAAAAAAAAAAAA 1816
Db      10090 GTTCAAGGCCAGCTGGGCACTAACAGATCTCTCTCAAAAAAAAAAAAAAAAAAAAA 10149
Qy      1817 AAGAAAGAGAGAGGAGCGGAGCTGTGGCTCACGCCGTATCCGACGACTTTGGAGGC 1876
Db      10150 AAGAAAGAGAGAGGAGCGGAGCTGTGGCTCACGCCGTATCCGACGACTTTGGAGGC 10209
Qy      1877 CGAGCCGGGCGGATCACTGTGTGTCAGAGTTTGAACCAAGCTGGCCAACTGGCAAA 1936
Db      10210 CGAGCCGGGCGGATCACTGTGTGTCAGAGTTTGAACCAAGCTGGCCAACTGGCAAA 10269
Qy      1937 CCCCCTGTGTACTCAAAATTCGCAAGCGGTGTAGCAGGCACTGTATCC 1996
Db      10270 CCCCCTGTGTACTCAAAATTCGCAAGCGGTGTAGCAGGCACTGTATCC 10329
Qy      1997 AGCTACTGGGAGGCTGAGCAGAGATGCTTGAACCAAGAGGTGAGGTTGAGTA 2056
Db      10330 AGCTACTGGGAGGCTGAGCAGAGATGCTTGAACCAAGAGGTGAGGTTGAGTA 10389
Qy      2057 AGCTGAGATGCTGCGCTTGCACTCCAGCTGGGCGAACAAGAGAGACTGTGTGAGA 2116
Db      10390 AGCTGAGATGCTGCGCTTGCACTCCAGCTGGGCGAACAAGAGAGACTGTGTGAGA 10449
Qy      2117 AAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2176
Db      10450 AAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 10509
Qy      2177 AAGCTTGCAAGAAATGTGCTTTATCCAACTAATGTAGAGCCAAATAGGATCCC 2236
Db      10510 AAGCTTGCAAGAAATGTGCTTTATCCAACTAATGTAGAGCCAAATAGGATCCC 10569
Qy      2237 TATTTGCTCTTTTGGTGTATTTGGTCCCTAACACTGTCTTTGCACTGGAAAAAA 2296
Db      10570 TATTTGCTCTTTTGGTGTATTTGGTCCCTAACACTGTCTTTGCACTGGAAAAAA 10629
Qy      2297 TTGAGATTAACATATCCCTGTCCTTATTAACCTGCAACCTTGCAATGAAGTAC 2356
Db      10630 TTGAGATTAACATATCCCTGTCCTTATTAACCTGCAACCTTGCAATGAAGTAC 10689
Qy      2357 AAGTCCACAGAAACTTGAATGACACACTGTCTTATTTAATCTTATTTGAATAGT 2416
Db      10690 AAGTCCACAGAAACTTGAATGACACACTGTCTTATTTAATCTTATTTGAATAGT 10749
Qy      2417 TGTAAAAAGATTAAATTTGTACTTCAATGATTTATTTATTTATTTATTTTGGCT 2476
Db      10750 TGTAAAAAGATTAAATTTGTACTTCAATGATTTATTTATTTATTTATTTTGGCT 10809
Qy      2477 CTAATGATTTTATTAACAGATTCCTTTCTGATATATTAATGAATGAAGTCAAGC 2536
Db      10810 CTAATGATTTTATTAACAGATTCCTTTCTGATATATTAATGAATGAAGTCAAGC 10869
Qy      2537 TTCAATTAATTTATTAATTTGAATGATTTCTAATAACAAGTATGTAATTTGACATTC 2596
Db      10870 TTCAATTAATTTATTAATTTGAATGATTTCTAATAACAAGTATGTAATTTGACATTC 10929
Qy      2597 AGTAATGTGCTACGAAGCAATTCCTTGAATTTAGTAATCTTTATGACAGCAAT 2656
Db      10930 AGTAATGTGCTACGAAGCAATTCCTTGAATTTAGTAATCTTTATGACAGCAAT 10989
Qy      2657 TGCTTCGGCTCACTTCAATCAGTTAAATTAATGAATTAATTTTGAAGTGTGAG 2716
Db      10990 TGCTTCGGCTCACTTCAATCAGTTAAATTAATGAATTAATTTTGAAGTGTGAG 11049
Qy      2717 ATAAATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2776
Db      11050 ATAAATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11109
Qy      2777 TATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11109

```

```

Db      11110 TATGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11130
RESULT 7
AAC55319
ID AAC55319 standard; DNA; 2172 BP.
XX
XX AAC55319;
AC
XX 05-FEB-2001 (first entry)
DT
XX
XX Human activation-induced cytidine deaminase exon 5 SRQ ID NO:15.
DE
XX
XX Activation-induced cytidine deaminase; AID; cytidine deaminase;
KM immune related disease; allergy; allergic disease; antiallergic;
KM antianemic; antiaesthetic; ophthalmological; anti-HIV; dermatological;
KM gene therapy; B cell associated immune system disorder; food allergy;
KM immunodeficiency disease; immunoglobulin A deficiency disease; ashrma;
KM IGA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis;
KM drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS;
KM ataxia telangiectasia; common variable immunodeficiency disorder;
KM major histocompatibility class II deficiency disease;
KM auto immunodeficiency syndrome; IgG subclass selection disorder; ds.
XX
CS Homo sapiens.
XX
XX WO200058480-A1.
PN
XX 05-OCT-2000.
PD
XX
XX 28-MAR-2000; 2000WO-JP001918.
PF
XX
XX 29-MAR-1999; 99JP-00087192.
PR 24-JUN-1999; 99JP-00178999.
PR 27-DEC-1999; 99JP-00371382.
XX
XX (NLSB ) JAPAN TOBACCO INC.
PA (HONU/) HONJO T.
PI
XX
XX Honjo T, Muramatsu M;
DR WPI; 2000-611715/58.
XX
XX Nucleic acid encoding activation induced cytidine deaminase, useful as a
PT target for drug development for immune-related diseases including
PT allergies.
XX
XX Claim 18; Page 152-153; 174pp; Japanese.
XX
XX The present invention describes an activation-induced cytidine deaminase
CC (AID). AID structurally relates to an RNA editing enzyme APOBEC-1 and has
CC cytidine activity similar to APOBEC-1. AID has antiallergic, antianemic,
CC antiaesthetic, ophthalmological, anti-HIV and dermatological activities,
CC and can be used in gene therapy. AID polynucleotides are useful in
CC methods for identifying drugs for the treatment of B cell associated
CC immune system disorder, immunodeficiency diseases and allergies, such as
CC immunoglobulin A (IGA) deficiency disease, IGA nephritis, gamma-
CC globulinaemia, atopic dermatitis, allergic colitis, asthma, food allergy,
CC drug allergy, allergic rhinitis, Rosen disease, Digorge disease, ataxia
CC telangiectasia, common variable immunodeficiency disorder, MHC (major
CC histocompatibility class II deficiency disease, AIDS (auto
CC immunodeficiency syndrome), elevated IGE disorder, and IGG subclass
CC selection disorder. The DNA sequences encoding AID may be used for gene
CC therapy and the antibodies to the AID protein may be used for diagnosis
CC and treatment of these disorders. The present sequence represents the
CC exon 5 genomic DNA sequence of human AID
XX
XX Sequence 2172 BP; 702 A; 379 C; 465 G; 626 T; 0 U; 0 Other;
Query Match 77.1%; Score 2172; DB 3; Length 2172;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 623 CCCCTGATAGGTTGATGACTTACAGACGCACTTTCGACTTTGGGACTTTGATAGCAA 682  
Db 1 CCCCTGATAGGTTGATGACTTACAGACGCACTTTCGACTTTGGGACTTTGATAGCAA 60  
QY 683 CTTCACAGAAATGTCACACAGATGAATAATCTCTGCTGAAGACAGTGAATAAAAAACAGT 742  
Db 61 CTTCACAGAAATGTCACACAGATGAATAATCTCTGCTGAAGACAGTGAATAAAAAACAGT 120  
QY 743 CCTTCAGTCTCTCTGTTTATTTTATCTTCACTCTCACTTTCTTAAAGTTTACAGAAAA 802  
Db 121 CCTTCAGTCTCTCTGTTTATTTTATCTTCACTCTCACTTTCTTAAAGTTTACAGAAAA 180  
QY 803 ATATTATATACAGACTCTTAAAGAAAGTCTATGCTTGAAGAAATAGAGAGAACAGAGT 862  
Db 181 ATATTATATACAGACTCTTAAAGAAAGTCTATGCTTGAAGAAATAGAGAGAACAGAGT 240  
QY 863 CTGGCCAGGAGCGTGTGCAATTTGTGCAAGTTTGAATGCAACATTTGCCCTACTGGGA 922  
Db 241 CTGGCCAGGAGCGTGTGCAATTTGTGCAAGTTTGAATGCAACATTTGCCCTACTGGGA 300  
QY 923 ATTAACAGACTGCAAGACCTGGAGACAATCTTAAGTCAACGTTTCTTATGACTTTTA 982  
Db 301 ATTAACAGACTGCAAGACCTGGAGACAATCTTAAGTCAACGTTTCTTATGACTTTTA 360  
QY 983 GGTAGATGAGAGAGAGTAGATCCTAAAGAAAGCAATGGTGAAGAGTCAATGTTTTA 1042  
Db 361 GGTAGATGAGAGAGAGTAGATCCTAAAGAAAGCAATGGTGAAGAGTCAATGTTTTA 420  
QY 1043 TATCAACATCCTTATATTTATTTGATTCATTTGATTAACAGTGGTGTATGATGATTTT 1102  
Db 421 TATCAACATCCTTATATTTATTTGATTCATTTGATTAACAGTGGTGTATGATGATTTT 480  
QY 1103 TCTATCTTTTCCCTTGAAGTTTACTTCAAGTAAACAACACTCTTCCATCAGGCCATGA 1162  
Db 481 TCTATCTTTTCCCTTGAAGTTTACTTCAAGTAAACAACACTCTTCCATCAGGCCATGA 540  
QY 1163 TCTATAGACCTCCTTAATGAGATCTGGGTATGATGACCCCAACCAATCTTCCCAA 1222  
Db 541 TCTATAGACCTCCTTAATGAGATCTGGGTATGATGACCCCAACCAATCTTCCCAA 600  
QY 1223 GCATTATATCCCATCATGCGCTGTATGTTTATGACGAGAGACAGTGTATGTTT 1282  
Db 601 GCATTATATCCCATCATGCGCTGTATGTTTATGACGAGAGACAGTGTATGTTT 660  
QY 1283 TACAAAAGAAATGTTATGGGTGGGATGAGATTAAGCAATGATGATCACTTCAA 1342  
Db 661 TACAAAAGAAATGTTATGGGTGGGATGAGATTAAGCAATGATGATCACTTCAA 720  
QY 1343 GCTACTTATTAAGAAATCTTAAATGGGAGAGACGTGAACAAGACCTTAATA 1402  
Db 721 GCTACTTATTAAGAAATCTTAAATGGGAGAGACGTGAACAAGACCTTAATA 780  
QY 1403 TGGGTGATGCTGAGATGACAAATCTTGGAAAGCAACCTTTTAAAGAGTCCCT 1462  
Db 781 TGGGTGATGCTGAGATGACAAATCTTGGAAAGCAACCTTTTAAAGAGTCCCT 840  
QY 1463 AATTAGAAACACCAACAACTTCAATATCATATAGCAACAAATGGAGAGAGTTG 1522  
Db 841 AATTAGAAACACCAACAACTTCAATATCATATAGCAACAAATGGAGAGAGTTG 900  
QY 1523 CTGTAATGTTGGGAGAGAAATCTATTTGGCTCTGTTGCTCTTCACTCAGAAATG 1582  
Db 901 CTGTAATGTTGGGAGAGAAATCTATTTGGCTCTGTTGCTCTTCACTCAGAAATG 960  
QY 1583 CCAATCAGTCAAGTCTTGTACATTTTGTATGCTGTGATGCTTCCCAAGATATAT 1642  
Db 961 CCAATCAGTCAAGTCTTGTACATTTTGTATGCTGTGATGCTTCCCAAGATATAT 1020  
QY 1643 TAACTATATAAGAGTTGTGACAAAACAGATGATTAAGCTGCAACCGTGCACACGC 1702  
Db 1021 TAACTATATAAGAGTTGTGACAAAACAGATGATTAAGCTGCAACCGTGCACACGC 1080  
QY 1703 TCATATGTTCTAGCTGTTGGAGAGTTGAGAGAGAGATGGCTTGAACACAGGTGTTCAA 1762

Db 1081 TCATATGTTCTAGCTGTTGGAGAGTTGAGAGAGAGATGGCTTGAACACAGGTGTTCAA 1140  
QY 1763 GGCACGCTGGGCAACATTAACAGATCTCTCTCAAAAAAAGAAAAAAGAAA 1822  
Db 1141 GGCACGCTGGGCAACATTAACAGATCTCTCTCAAAAAAAGAAAAAAGAAA 1200  
QY 1823 GAGAGAGAGGCGGGCGGTGTGGCTCAAGCTGTAAATCCAGCACTTTGGAGGCGAGGC 1882  
Db 1201 GAGAGAGAGGCGGGCGGTGTGGCTCAAGCTGTAAATCCAGCACTTTGGAGGCGAGGC 1260  
QY 1883 GAGCGGATCACTGTGTGTGAGAGTTTGAAGCCAGCTGTGCAACATGGCAAAACCCGT 1942  
Db 1261 GAGCGGATCACTGTGTGTGAGAGTTTGAAGCCAGCTGTGCAACATGGCAAAACCCGT 1320  
QY 1943 CTGTACTCAAAATGCAAAAAATTAACAGCGGTGTGAGAGCACTGTGTAATCCAGTAC 2002  
Db 1321 CTGTACTCAAAATGCAAAAAATTAACAGCGGTGTGAGAGCACTGTGTAATCCAGTAC 1380  
QY 2003 TTGGAGAGCTGAGGACAGAGAAATGCTTGAATCCAGAGAGTGAAGTTGAGTGA 2062  
Db 1381 TTGGAGAGCTGAGGACAGAGAAATGCTTGAATCCAGAGAGTGAAGTTGAGTGA 1440  
QY 2063 GATCGTCCGTTGCACTCCAGCTGGGCGCAAGAGCAAGTCTGTCTCAGAAAAA 2122  
Db 1441 GATCGTCCGTTGCACTCCAGCTGGGCGCAAGAGCAAGTCTGTCTCAGAAAAA 1500  
QY 2123 AAAAAAAGAT 2182  
Db 1501 AAAAAAAGAT 1560  
QY 2183 TGCAGAGAAATTTGCTTATTCACAAATGTAAGAGCAATTAAGGATCCCTATTTG 2242  
Db 1561 TGCAGAGAAATTTGCTTATTCACAAATGTAAGAGCAATTAAGGATCCCTATTTG 1620  
QY 2243 TCTCTTTTGTGTATTTTGTCCCTAACATCTGTCTGACAGTGAAGAAATATTCAGA 2302  
Db 1621 TCTCTTTTGTGTATTTTGTCCCTAACATCTGTCTGACAGTGAAGAAATATTCAGA 1680  
QY 2303 ATTAACATATCCCTGTGCGCTTATTAACCTAGCAACCTTGCATTAAGATGAGAGATCC 2362  
Db 1681 ATTAACATATCCCTGTGCGCTTATTAACCTAGCAACCTTGCATTAAGATGAGAGATCC 1740  
QY 2363 AAGAGAAACTTGAATGACACGTCTTATTTTATTTATTTATTTATTTATTTATTTATTT 2422  
Db 1741 AAGAGAAACTTGAATGACACGTCTTATTTTATTTTATTTATTTATTTATTTATTTATTT 1800  
QY 2423 AGAGTAAAAATTTGTACTTCAATGATTAATTTATTTATTTATTTATTTATTTATTTATTT 2482  
Db 1801 AGAGTAAAAATTTGTACTTCAATGATTAATTTATTTATTTATTTATTTATTTATTTATTT 1860  
QY 2483 ATTTTTATTAACATGATTTCTTTCTGTATTAATTAAGAAATGAGTCTCAAGCTTCATA 2542  
Db 1861 ATTTTTATTAACATGATTTCTTTCTGTATTAATTAAGAAATGAGTCTCAAGCTTCATA 1920  
QY 2543 AATTATTAACCTTGAAGATGATCTTAATAACAAGTATGTAATTTGAACATTTGAGATAT 2602  
Db 1921 AATTATTAACCTTGAAGATGATCTTAATAACAAGTATGTAATTTGAACATTTGAGATAT 1980  
QY 2603 GGTGTACAGAGCAATTTCTTGTATTTTATTAAGTATTTTATTAAGAGCAATTTGCTTC 2662  
Db 1981 GGTGTACAGAGCAATTTCTTGTATTTTATTAAGTATTTTATTAAGAGCAATTTGCTTC 2040  
QY 2663 TGGCTCACTTCAATCAGTTTAAATAAGTAAATTAATTTTGAAGCTGAGAGATATAA 2722  
Db 2041 TGGCTCACTTCAATCAGTTTAAATAAGTAAATTAATTTTGAAGCTGAGAGATATAA 2100  
QY 2723 TACCAATTAATAATTAATAAAGTATTAATGATTAATAAATTAATAATTAATAATTAATA 2782  
Db 2101 TACCAATTAATAATTAATAAAGTATTAATGATTAATAAATTAATAATTAATAATTAATA 2160  
QY 2783 GGAATTAACCTTG 2794

Db 2161 GGAATAACTTG 2172

## RESULT 8

ABX05468 standard; cDNA; 1543 BP.

AC ABX05468;

DT 17-JAN-2003 (first entry)

DE Human novel polynucleotide #483.

Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer; neurodegenerative disorder; lymphoid cell disorder; osteoporosis; Parkinson's disease; Alzheimer's disease; bone degenerative disorder; osteoarthritis; periodontal disease; liver fibrosis; viral infection; fungal infection; bacterial infection; autoimmune disease; diabetes; atopic dermatitis.

OS Homo sapiens.

PN WO200274961-A1.

PD 26-SEP-2002.

PF 14-MAR-2002; 2002WO-US005109.

PR 15-MAR-2001; 2001US-00810173.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F, Xue J, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M, Wehrman T, Wang J, Wang D, Drmanac RT;

PI WPI; 2003-040556/03.

DR P-PSDB; ABU00390.

DR New isolated polypeptides and polynucleotides, useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders, lymphoid cell disorders, bone degenerative disorders, and infections.

PT Claim 1; SEQ ID NO 483; 235bp; English.

PS The invention relates to human polynucleotides and the polypeptides they encode. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, medical imaging, identification of mutations, responsible for genetic disorders or other traits, assessing biodiversity and producing many other types of data and products dependent on DNA and amino acid sequences. They are also useful for preventing, treating or ameliorating medical conditions, such as cancer, neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell disorders, osteoporosis, osteoarthritis, bone degenerative disorders, periodontal disease, liver fibrosis, infections (e.g. viral, fungal or bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).

CC Sequences ABX04986-ABX0551 represent human polynucleotides of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied by the European Patent Office

CC Sequence 1543 BP; 428 A; 334 C; 352 G; 429 T; 0 U; 0 Other;

SQ Query Match 54.7%; Score 1542; DB 7; Length 1543;

Best Local Similarity 100.0%; Pred. No. 1.2e-258;

Matches 1542; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAAACATCAATTAATGAGAGATTTCTGCGCTGAGACTGCGAGGAGGAGAGAA 63

Db 2 GAAACATCAATTAATGAGAGATTTCTGCGCTGAGACTGCGAGGAGGAGAGAA 61

QY 64 CACTGTGACACACTATGAGACGCTCTTGATGAACCGAGGAAAGTTCTTACCAATT 123

Db 62 CACTGTGACACACTATGAGACGCTCTTGATGAACCGAGGAAAGTTCTTACCAATT 121

QY 124 CAAAATGTCGCTGAGGCTAAGGCTGAGACCTACCTGTGCTACGATGAAAG 183

Db 122 CAAAATGTCGCTGAGGCTAAGGCTGAGACCTACCTGTGCTACGATGAAAG 181

QY 184 GCGTGAACGCTACATCCCTTTTCACTGAGATTGGTATTTCTTGGCAATAGAA 243

Db 182 GCGTGAACGCTACATCCCTTTTCACTGAGATTGGTATTTCTTGGCAATAGAA 241

QY 244 CCACTGTGAATGCTCTCTCCGCTACATCTGCACTGAGACCTAGACCTGAGCGCTG 303

Db 242 CCACTGTGAATGCTCTCTCCGCTACATCTGCACTGAGACCTAGACCTGAGCGCTG 301

QY 304 CTACGCGCTACCTGCTTACCTCTGAGAGCCCTGCTACGACTGTGACCATGAGC 363

Db 302 CTACGCGCTACCTGCTTACCTCTGAGAGCCCTGCTACGACTGTGACCATGAGC 361

QY 364 CCACTTTTCTGAGAGAAACCCCAACCTCAGCTGAGATCTTACCGCGGCTCTACTT 423

Db 362 CCACTTTTCTGAGAGAAACCCCAACCTCAGCTGAGATCTTACCGCGGCTCTACTT 421

QY 424 CTGTGAGAACCGGAGAGCTGAGCCCGAGGCTGCGGCTGACCGCGCGGCTGCA 483

Db 422 CTGTGAGAACCGGAGAGCTGAGCCCGAGGCTGCGGCTGACCGCGCGGCTGCA 481

QY 484 AATAGCCATCATGACCTTCAAGATTTTCTGCTGAGAAATCTTTGAGAAACCA 543

Db 482 AATAGCCATCATGACCTTCAAGATTTTCTGCTGAGAAATCTTTGAGAAACCA 541

QY 544 TGAAGAACTTTCAAGCTGAGAGGCTGACATGAAATCTGCTCTGACAGA 603

Db 542 TGAAGAACTTTCAAGCTGAGAGGCTGACATGAAATCTGCTCTGACAGA 601

QY 604 GCTTGGGCGCATCTTTTGGCCCTGATGAGCTTGAAGATGATGAGAGAGCATTTGCTAC 663

Db 602 GCTTGGGCGCATCTTTTGGCCCTGATGAGCTTGAAGATGATGAGAGAGCATTTGCTAC 661

QY 664 TTTGGGACTTTGATGAGAACTTCCAGAAATGTCACACAGATGAATATCTGCTGAG 723

Db 662 TTTGGGACTTTGATGAGAACTTCCAGAAATGTCACACAGATGAATATCTGCTGAG 721

QY 724 ACAGTGAATTAAGAAAGCTCTTCAAGCTTCTGTTTATCTTCAACTCTCACTT 783

Db 722 ACAGTGAATTAAGAAAGCTCTTCAAGCTTCTGTTTATCTTCAACTCTCACTT 781

QY 784 CTTGAGTTTACAGAAATATTTATATACAGCTCTTTAAAGATCTATGCTTGA 843

Db 782 CTTGAGTTTACAGAAATATTTATATACAGCTCTTTAAAGATCTATGCTTGA 841

QY 844 ATAGAGAAAGAAACAGCTCTGAGAGAGCTGAGAGCTGAGAGCTTGAAGTGA 903

Db 842 ATAGAGAAAGAAACAGCTCTGAGAGAGCTGAGAGCTGAGAGCTTGAAGTGA 901

QY 904 ACATTTGCTCTTCTGAGAAATTAACAGAACTGAGAGCTTGAAGTGAAGTGA 963

Db 902 ACATTTGCTCTTCTGAGAAATTAACAGAACTGAGAGCTTGAAGTGAAGTGA 961

QY 964 GCTTTTCTATGACTTTTGAAGATGAGAGAGAGATGATCTTAAAGAGATGAGT 1023

Db 962 GCTTTTCTATGACTTTTGAAGATGAGAGAGAGATGATCTTAAAGAGATGAGT 1021

QY 1024 AGAGATCAATGTTTATATCAACATCTTATATTTGATTCATTTGATTAAGT 1083

Db 1022 AGAGATCAATGTTTATATCAACATCTTATATTTGATTCATTTGATTAAGT 1081

QY 1084 GGTGTATGATTAATTTTCTATCTTTTCCCTGAGCTTCACTTCAAGTAAACAA 1143

Db 1082 GGTGTATGATTAATTTTCTATCTTTTCCCTGAGCTTCACTTCAAGTAAACAA 1141

QY 1144 CTTTTCATCAAGCCATGATCTATAGACCTCTTATGAGATGATGAGTGGTGTGATGAC 1203





PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 05-DEC-2000; 2000US-0250306P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM,  
 XX MPI; 2001-483426/52.  
 DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
 XX  
 XX  
 PS Disclosure; SEQ ID NO 35900; 3071pp + Sequence Listing; English.  
 XX  
 CC AAK5951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytosolic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK81694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169  
 CC represent sequences used in the exemplification of the present invention  
 XX  
 XX Sequence 1665 BP; 497 A; 343 C; 365 G; 460 T; 0 U; 0 Other;

QY 617 CTTTGGCCCTGATGAGTGTGATGACTACGAGCGACTTTCGTACTTTGGGACTTTGA 676  
 Db 172 CTACAGCCCTGTATGAGGTGATGATGACTACGAGCGACTTTCGTACTTTGGGACTTTGA 231  
 QY 677 TAGCACTTCAGGAATGTGACACAGATGAAATCTCTGCTGAAGACATGATGATAA 736  
 Db 232 TAGCACTTCAGGAATGTGACACAGATGAAATCTCTGCTGAAGACATGATGATAA 291  
 QY 737 AACAGCTTCAGGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 796  
 Db 292 AACAGCTTCAGGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 351  
 QY 797 GAAAAAATATTTATATAGACTCTTTAAAGATCATGTCTTGAATAGAGAGAAC 856  
 Db 352 GAAAAAATATTTATATAGACTCTTTAAAGATCATGTCTTGAATAGAGAGAAC 411  
 QY 857 ACAGGCTGCGCCAGGAGCGTCTGCAATTGTCAGCTTTGAAATGCAATGCTCCCTA 916  
 Db 412 ACAGGCTGCGCCAGGAGCGTCTGCAATTGTCAGCTTTGAAATGCAATGCTCCCTA 471  
 QY 917 CTGGGAATACAGAACTGCAAGACTGCGAGCACTCTTAAAGTGTCAAGCTTTTCTATGA 976  
 Db 472 CTGGGAATACAGAACTGCAAGACTGCGAGCACTCTTAAAGTGTCAAGCTTTTCTATGA 531  
 QY 977 CTTTGAAGTAGATGAGACAGAGAGTAGATCTTAAAGCATGCTGAGAGAGATCAATG 1036  
 Db 532 CTTTGAAGTAGATGAGACAGAGAGTAGATCTTAAAGCATGCTGAGAGAGATCAATG 591  
 QY 1037 TTTTATATCAACATCTTTATATTTGATTCATTTGATTAACAGTGTGTTAGATATA 1096  
 Db 592 TTTTATATCAACATCTTTATATTTGATTCATTTGATTAACAGTGTGTTAGATATA 651  
 QY 1097 GATTTTCTATTTCTTTCCCTTGACCTTTCAGTTTCAAGTAACAACATCTTTCATCAG 1156  
 Db 652 GATTTTCTATTTCTTTCCCTTGACCTTTCAGTTTCAAGTAACAACATCTTTCATCAG 711  
 QY 1157 CCATGATCTATAGAGACTCTCTAATGAGATCTCTGCTGATTTGTGACCCCAACCATCTC 1216  
 Db 712 CCATGATCTATAGAGACTCTCTAATGAGATCTCTGCTGATTTGTGACCCCAACCATCTC 771  
 QY 1217 TCCAAAGCATTAATATCCATCATGCGCTGTATGTTTATATAGCAAGACATGTTTGA 1276  
 Db 772 TCCAAAGCATTAATATCCATCATGCGCTGTATGTTTATATAGCAAGACATGTTTGA 831  
 QY 1277 TGTTTGACAAAGAGATTTGATGAGTGGGATGAGATGAGATGAGATGAGATGAGATG 1336  
 Db 832 TGTTTGACAAAGAGATTTGATGAGTGGGATGAGATGAGATGAGATGAGATGAGATG 891  
 QY 1337 CTTCAGGCTATCTTTAATTAAGATCTTAAATGGGCGAGAGACTGTGAACAAGACACC 1396  
 Db 892 CTTCAGGCTATCTTTAATTAAGATCTTAAATGGGCGAGAGACTGTGAACAAGACACC 951  
 QY 1397 TAATTAATGGGTTGATGTCGAGTGAAGTCAATCTCTGGAAGCAACCTTTTAAAGAA 1456  
 Db 952 TAATTAATGGGTTGATGTCGAGTGAAGTCAATCTCTGGAAGCAACCTTTTAAAGAA 1011  
 QY 1457 GTCCCTAATTTAGAAACCCCAACCACTTCAATCATATATATAGCAACCAATTTGAGAG 1516  
 Db 1012 GTCCCTAATTTAGAAACCCCAACCACTTCAATCATATATATAGCAACCAATTTGAGAG 1071  
 QY 1517 AAGTTCCTGATGTTGGGAGAGAGAAATCTATTTGCTCTGCTGCTCTTTCATCTCA 1576  
 Db 1072 AAGTTCCTGATGTTGGGAGAGAGAAATCTATTTGCTCTGCTGCTCTTTCATCTCA 1131  
 QY 1577 GAATTCGCAATGAGTGTGAGAGTGTGATATTTGATGATGATGATGATGATGATGATG 1636  
 Db 1132 GAATTCGCAATGAGTGTGAGAGTGTGATATTTGATGATGATGATGATGATGATGATG 1191  
 QY 1637 GTATATTAATATATATAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1696  
 Db 1192 GTATATTAATATATATAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1251

ID	Accession	Standard	CDNA	BP
QY	1697	ACACGCTCATATGATCTAGATCGTGTGGAGGTTTGAGAGGAGAGATGGCTTGAAACAAGT	1756	
Db	1252	ACACGCTCATATGATCTAGATCGTGTGGAGGTTTGAGAGGAGAGATGGCTTGAAACAAGT	1311	
QY	1757	GTTTCAAGGCGAGCTTGAGGCAATTAACAAGATCTCTGTCTCTCAAAAAA	1816	
Db	1312	GTTTCAAGGCGAGCTTGAGGCAATTAACAAGATCTCTGTCTCTCAAAAAA	1377	
QY	1817	AAGAAAGAGAGAGGCGCGGAGCTGGTGGCTCAAGCTCTGTATTCACACTTTGGAGGC	1876	
Db	1372	AAGAAAGAGAGAGGCGCGGAGCTGGTGGCTCAAGCTCTGTATTCACACTTTGGAGGC	1433	
QY	1877	CGAGCCGGGGGATCACTGTGTGTGAAGAGTTTGAAGACAGCCTGGCAATGGCAAA	1933	
Db	1432	CGAGCCGGGGGATCACTGTGTGTGAAGAGTTTGAAGACAGCCTGGCAATGGCAAA	1491	
QY	1937	CCCCGTCTGTACTCAAAATGCAAAATTAAGCAGGCGTGGTAGAGCACTGTATCCC	1996	
Db	1492	CCCCGTCTGTACTCAAAATGCAAAATTAAGCAGGCGTGGTAGAGCACTGTATCCC	1551	
QY	1997	AGTACTTGGGAGGCTGAGGCAAGGAAATGCTTGAACCAAGAGGTGAAGGTTGACATTA	2056	
Db	1552	AGTACTTGGGAGGCTGAGGCAAGGAAATGCTTGAACCAAGAGGTGAAGGTTGACATTA	1611	
QY	2057	AGCTGAGATCGGCGGTTGCAATCCAGGCTGGGCGCAAGACAAATCTGTGC	2110	
Db	1612	AGCTGAGATCGGCGGTTGCAATCCAGGCTGGGCGCAAGACAAATCTGTGC	1665	
RESULT 10				
ID	AAAC5307	standard	CDNA	2440 BP
XX	AAAC5307			
XX	AAAC5307			
DT	05-FEB-2001	(first entry)		
XX				
DE	Mouse activation-induced cytidine deaminase encoding cDNA SEQ ID NO:1.			
XX				
KW	Activation-induced cytidine deaminase; AID; cytidine deaminase; immune related disease; allergy; allergic disease; antiallergic; antianemic; antisthmatic; ophthalmological; anti-HIV; dermatological; gene therapy; B cell associated immune system disorder; food allergy; immunodeficiency disease; immunoglobulin A deficiency disease; asthma; IgA nephritis; gamma-globulinaemia; atopic dermatitis; allergic colitis; drug allergy; allergic rhinitis; Rosen disease; Digorge disease; AIDS; ataxia telangiectasia; common variable immunodeficiency disorder; major histocompatibility class II deficiency disease;			
KW	auto immunodeficiency syndrome; IgG subclass selection disorder; ss.			
XX				
OS	Mus musculus.			
XX				
FH	Location/Qualifiers			
FT	CDS	93..689		
FT		/tag= a		
FT		/product= "activation-induced cytidine deaminase"		
XX				
PN	WO200058480-A1.			
XX				
PD	05-OCT-2000.			
XX				
PF	28-MAR-2000; 2000MO-JP001918.			
XX				
PR	29-MAR-1999; 99JP-00087192.			
PR	24-JUN-1999; 99JP-00178989.			
PR	27-DEC-1999; 99JP-00371382.			
XX				
PA	(NIBS ) JAPAN TOBACCO INC.			
PA	(HONTU/) HONTO T.			
XX				
XX	Honjo T, Muramatsu M,			
XX	WPI; 2000-611715/58.			

Query Match	21.4%	Score 603.4	DB 3	Length 2440
Best Local Similarity	69.4%	Pred. No. 1.4e-95		
Matches 1006	Conservative 0	Mismatches 351	Indels 92	Gaps 10
Sequence 2440 BP	706 A	546 C	551 G	636 T
	0 U	1 Other		
Query	4	GAACCATCATTAATTAAGAGAGATTTTCTGCGCTGAGACTTGCAGGAGAGCGAAG-AGG	62	
Db	16	GAAGCAGCGCTTGCTTGAAGCAGGCTTCCTTTGGCTTAAGACTTTGAGGAGTCAAGAAAG	75	
Query	63	ACACCTTGGACACACACTATGAGACGCTCTTGATGAACCGAGAGAGTTCTTTACCAT	122	
Db	76	TACAGCTGAGAGCCGATATGACAGGCTTCTGATGAAGCAAAAGAGTTTCTTTACCAT	135	
Query	123	TCAAAATATCCGCTGGGCTTAAGGCTGAGGCTGAGACCTACCTGCTGACGTAGTAAAG	182	
Db	136	TCAAAATATCCGCTGGGCTGAGGCTGAGGCTGAGACCTACCTGCTGACGTAGTAAAG	195	
Query	183	GCGGTGACAGTCTCATCTCTTTCACATGAGCTTTGGTTATCTTTCGCAATTAAGACGGCT	242	
Db	196	GAGAGATAGTCCACCTCTGCTGACATCTGAGCTTGGCCACCTTGGCAACAGTCCGGCT	255	
Query	243	GCCACGTGAATGCTCTCTTCTCCGCTACATCTGAGCTGGAGCTTAAGACCTGGCGCT	302	
Db	256	GCCACGTGAATGCTCTCTTCTCCGCTACATCTGAGCTGGAGCTTAAGACCTGGCGCT	315	
Query	303	GCTACCGGCTCACTGCTGTTCACTCTCGAGGCGCTGCTGAGCTGAGCTGAGCAATGTGG	362	
Db	316	GTTTACCGCTCACTGCTGTTCACTCTCGAGGCGCTGAGCTGAGCTGAGCAATGTGG	375	
Query	363	CCGACTTCTGCGAGGAGAACCCCAACTCACTGCTGAGATCTTACCCGCGCTCTACT	422	
Db	376	CTGAGTTTCTGAGATGAGAACCTTAACCTCACTGAGCTGAGATTTTCAACCGCGCTCTACT	435	
Query	423	TCTGTAGGACCGCAAGGCTAGGCCGAGGGGCTGCGGCGGCTGACCCGCGCGGGGTGC	482	
Db	436	TCTGTAGGACCGCAAGGCTAGGCCGAGGGGCTGCGGCGGCTGACCCGCGCGGGGTGC	495	
Query	483	AAATACCATCATGACCTTCAAAAGTATTTTATCTGCTGGGAATCACTTTTGTAGAAAC	542	
Db	496	AGATCGGAGTCAATGACCTTCAAAAGTATTTTATCTGCTGGGAATCACTTTTGTAGAAAC	555	
Query	543	ATGAAGAAGCTTCAAAAGCTGAGGAGGCTGATGAATAATGATGTTGCTCTCCAGAC	602	
Db	556	GTTAAAGAAGCTTCAAAAGCTGAGGAGGCTGATGAATAATGATGTTGCTCTCCAGAC	615	
Query	603	AGCTTGGGAGATCTTTTGGCCCTGTATGAGTTATGATCTTACGAGAGCGATTTTCGTA	662	

Db	616	AACTTGAGGAGCATCCTTTTGCCCTTGTAACAAATGCATGACTTGCAGAGATGCATTTCCGTA	675
Qy	663	CTTTGGGACCTTTGATATGCAACTTCCAGGAATGTCAACAAGATGAATAATCTGTGTGAA	722
Db	676	TGTTGGGATTTTGAAAGCAACCTCCGGAATGTCAACGATGAATAATTTCT---CTGAA	732
Qy	723	GACAGTGAATTAATAACAAGCTCTTCAGAGTCTCTGTGTTTTATCTTCAACTCTCACTT	782
Db	733	GAGACTGAAATGAAAAACAACCTTCA--ACTACATGTTTTCTTTTAAGTACTCACTT	790
Qy	783	TCCTAAGTTTACAGAAAAAATATTTATATACACTCTTTAAAAAGATCTAGCTTGA	842
Db	791	TTATATAGTAGGGGAAA-----TTATATACCTTTTAAAAAATCTTGAGCTGCAC	843
Qy	843	AATAGAAAGAAACACAGGTCTGGCCAGGACGTGTGCATTTGTGCAGTTTGTGATGC	902
Db	844	AGGACCGCCACAGACATGATGTAC-----TGAGCTTGCCTGTC	882
Qy	903	AACATTGTGTCCTACT-GGGAATACAGAACTGCAGACCTGGAGCATCTTAAGTGTG	961
Db	883	AACATGCCATCTACCTGGGAAACAGATTAATCCAGCTTTGGGTGCGAATGATGCTC	942
Qy	962	AACGTTTTCTATGACTTTTAGTAGAGTAGAGAGAGAAAGTATGATCCATAAAGATGG	102
Db	943	TTTTTTTCAAGCAT-----GAAAAGCATAT	971
Qy	1022	TGAGAGATCAATGTTTTTATATCAACATCCCTTATATTTGATTCATTTGAGTTAACA	108
Db	972	GGAGAGACACACAGATTTGTAAACCAACCGTGTTCTTGATTTCAATTTGAATTTCTCA	103
Qy	1082	GTGGTGTATGTATAGATTTTCTATCTTTTCCCTTGACGTTTACTTCAAGTAAACA	114
Db	1032	GGGATATCAGTAGCGATTTCTTATCTTCCCTTAAGGCTCACTTCAGG-----	108
Qy	1142	AACTCTTCATGAGGCGCATGATCTATAGGACGCTCCATAGAGATCTGGGTGATTTGG	120
Db	1085	---GGTCTTTTCTGACAAAGTCAAGGAGCTGTCTCAAGCTCTGTGTAGCAAT---	113
Qy	1202	ACCCCAACCATCTCTCCAAAGCATTAATATCCAAATCATAGCGCTGATGTGTTAATCAGC	126
Db	1138	-CACAAAGCATCTCTCAAAAAATTAAATCAAGGCAATCTGTATGTTC-----	119
Qy	1262	AGAAGCATGTTTTATGTTTGTACAAAGAATGTTATGSGTGGGATGAGGATNAG	132
Db	1192	-----CTGTCCGTGCTGTTTTTCACTTTGTATGTAGAAAGGGCTTGGGATTTTGA	124
Qy	1322	ACCATGACATGATGACCTTCAAGCTACTTTAATAAAGATCTTAATAATGGCAGAGACT	138
Db	1246	AGAAATGACAGATGCGCTCTGGGTGATTTCAATAAAGAGATCTTAATAATGAGATGAGACT	130
Qy	1382	GTGAACAAGACACCTTAATATAGGTGTGATGTCTAAGTAGCAAAATCTTGTGAACGCA	144
Db	1306	ACGAAGAATCACTCTGAAAAAGATTCAGCGCTCAAGAAGCAATCCCTGTGAACACA	136
Qy	1442	AACTCTTTT 1450	
Db	1366	GACTCTTTT 1374	
RESULT 11			
ADBB88952 standard; DNA; 597 BP.			
ADBB88952;			
04-DEC-2003 (first entry)			
AID gene as substrate used in AID mutagenic method.			
ds; cytosstatic; virucide; antidiol; mutation;			
activation-induced cytidine deaminase; cell phenotype;			
monoclonal antibody; antigen; cross-reactivity; tumour;			
germ warfare agent; toxin; V region.			



RESULT 12  
AAK81089  
ID AAK81089 standard; DNA; 574 BP.  
XX  
XX AAK81089;  
AC  
XX  
DT 07-NOV-2001 (first entry)  
XX  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35901.  
DE  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW Cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216800P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 27-SEP-2000; 2000US-0235837P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236328P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249247P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259578P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX  
XX  
XX Disclosure; SEQ ID NO 35901; 3071bp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention  
XX  
XX Sequence 574 BP; 201 A; 80 C; 85 G; 208 T; 0 U; 0 Other;  
SQ  
Query Match 20.4%; Score 574; DB 4; Length 574;  
Best Local Similarity 100.0%; Pred. 1.5e-90;  
Matches 574; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2211 AATGTAAGAGCCAAATAGGAGATCCCTATTGCTCTTTGGTGTCTATTGCTCTTAAC 2270  
DB 1 AATGTAAGAGCCAAATAGGAGATCCCTATTGCTCTTTGGTGTCTATTGCTCTTAAC 60  
QY 2271 AACTGCTCTTGACGTGAGAAAAATATTGAGATTAACCATATCCCTGCGCTATTAC 2330  
DB 61 AACTGCTCTTGACGTGAGAAAAATATTGAGATTAACCATATCCCTGCGCTATTAC 120  
QY 2331 TAGCAACCCCTGCATGAGATGAGAGATCCACAGAAAATTGTAATGCAACTGCT 2390  
DB 121 TAGCAACCCCTGCATGAGATGAGAGATCCACAGAAAATTGTAATGCAACTGCT 180  
QY 2391 TATTTAATCTTATGATAGTAAAGTGTAAAGAGTAAAGTGTAAAGTGTAAAGTGTAA 2450  
DB 181 TATTTAATCTTATGATAGTAAAGTGTAAAGAGTAAAGTGTAAAGTGTAAAGTGTAA 240  
QY 2451 CATTTAATTTTATTTATTTGCTATGATTTTATTAACATGATTTCTTTCT 2510  
DB 241 CATTTAATTTTATTTATTTGCTATGATTTTATTAACATGATTTCTTTCT 300  
QY 2511 GATATATGAAATGAGATCAAGCTTCAATTAATTAATTAATTAATTAATTAATTAAT 2570  
DB 301 GATATATGAAATGAGATCAAGCTTCAATTAATTAATTAATTAATTAATTAATTAAT 360  
QY 2571 AACACGTATGATTAATTTAAATTCAGTATGATGCTAGCAAGCAATTTCTTGAATT 2630

DB 361 AACACGTATGATTAATTTAAATTCAGTATGATGCTAGCAAGCAATTTCTTGAATT 420  
QY 2631 TTAGTAAATCTTTATGACAGCAAAATTTGCTTGTGCTCACTTCAATCAGTTAAAT 2690  
DB 421 TTAGTAAATCTTTATGACAGCAAAATTTGCTTGTGCTCACTTCAATCAGTTAAAT 480  
QY 2691 GATATATATTTTGGAGCTGGAAGTAAATATCCAAATTAATTAATTAATTAATTAAT 2750  
DB 481 GATATATATTTTGGAGCTGGAAGTAAATATCCAAATTAATTAATTAATTAATTAAT 540  
QY 2751 TATATGAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2784  
DB 541 TATATGAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 574  
RESULT 13  
AAK61819  
ID AAK61819 standard; cDNA, 429 BP.  
XX  
XX AAK61819;  
XX  
XX 06-NOV-2001 (first entry)  
XX  
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6879.  
DE  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ss.  
XX  
XX Homo sapiens.  
PN WO200157182-A2.  
PD 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001354.  
PF  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186359P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225457P.  
PR 14-AUG-2000; 2000US-0225788P.  
PR 14-AUG-2000; 2000US-0225789P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226641P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.



PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230457P.  
PR 06-SEP-2000; 2000US-0230458P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235444P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236337P.  
PR 29-SEP-2000; 2000US-0236338P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241231P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241866P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250319P.  
PR 05-DEC-2000; 2000US-0251038P.  
PR 05-DEC-2000; 2000US-0251088P.  
PR 05-DEC-2000; 2000US-0251798P.  
PR 06-DEC-2000; 2000US-0256719P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254037P.  
PR 05-JAN-2001; 2001US-02595678P.  
  
(HOMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
MPI; 2001-483426/52.  
P-PSDB; AAM89038.  
  
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
useful for preventing, diagnosing and/or treating cancers and metastasis.  
  
Claim 1; SEQ ID NO 6879; 3071bp + Sequence Listing; English.  
  
AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK7694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
CC  
CC  
SQ Sequence 429 BP; 144 A; 79 C; 97 G; 109 T; 0 U; 0 Other;  
  
Query March 15.2%; Score 429; DB 4; Length 429;  
Best Local Similarity 100.0%; Pred. No. 2.2e-65;  
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1386 ACAAGACACCTATATATGAGTGTGAGTAGCAATCTTGTGAAAGCAAACT 1445  
DB 1 ACAAGACACCTATATATGAGTGTGAGTAGCAATCTTGTGAAAGCAAACT 60  
QY 1446 CTTTAAAGAGAGTCCCTAATTAGAAACACCAACCACTTACATATCATTAATAGCAA 1505  
DB 61 CTTTAAAGAGAGTCCCTAATTAGAAACACCAACCACTTACATATCATTAATAGCAA 120  
QY 1506 CATTGGAAGAGAGTCTGTAATGTGGGGAAGAGAAATCTAATGAGCTCTGTGGGTC 1565

Db 121 CAATTGGAAGAAATGCTTGAATGTTGGGAGAGAAAACTATTGCTCTGCTGGCTC 180  
Qy 1566 TCTTCATCTCAGAAATGCCATCAGTCAAGTGTGCTAGATTTGTATGTGTGATGC 1625  
Db 181 TCTTCATCTCAGAAATGCCATCAGTCAAGTGTGCTAGATTTGTATGTGTGATGC 240  
Qy 1626 TTTCTCCCAAGGTATTAATAATATAAGAGAGTTGACAAAACAGATGATAAAGCTG 1685  
Db 241 TTTCTCCCAAGGTATTAATAATATAAGAGAGTTGACAAAACAGATGATAAAGCTG 300  
Qy 1686 CGAACCGTGGACACGCTCATAGTCTAGCTGTGGAGAGTTGAGAGGAGAGATGCGCT 1745  
Db 301 CGAACCGTGGACACGCTCATAGTCTAGCTGTGGAGAGTTGAGAGGAGAGATGCGCT 360  
Qy 1746 TGAACACAGGTGTTCACAGGCGCCTGGGCAACATAACAGATCTGTCTCAAAAAA 1805  
Db 361 TGAACACAGGTGTTCACAGGCGCCTGGGCAACATAACAGATCTGTCTCAAAAAA 420  
Qy 1806 AAAAAAAAA 1814  
Db 421 AAAAAAAAA 429  
RESULT 14  
ID AAL02789 standard; DNA; 13862 BP.  
AC AAL02789;  
XX 21-NOV-2001 (first entry)  
DT Human reproductive system related antigen DNA SEQ ID NO: 5477.  
DE Human reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ds.  
KM  
XX Homo sapiens.  
OS  
PV W0200155320-A2.  
PD 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US001339.  
PF 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 26-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226682P.  
PR 22-AUG-2000; 2000US-0227183P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229347P.  
PR 01-SEP-2000; 2000US-0229348P.  
PR 01-SEP-2000; 2000US-0229349P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229515P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0233968P.  
PR 14-SEP-2000; 2000US-0233979P.  
PR 14-SEP-2000; 2000US-0233989P.  
PR 14-SEP-2000; 2000US-0233999P.  
PR 14-SEP-2000; 2000US-0234009P.  
PR 14-SEP-2000; 2000US-0234010P.  
PR 14-SEP-2000; 2000US-0234011P.  
PR 14-SEP-2000; 2000US-0234063P.  
PR 14-SEP-2000; 2000US-0234064P.  
PR 14-SEP-2000; 2000US-0234065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235833P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241212P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0246417P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.

```

PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251866P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-46570/50.
XX
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen is
XX
XX used in preventing, treating or ameliorating a medical condition.
XX
XX
PS Disclosure; SEQ ID NO 5477; 1297bp + Sequence Listing; English.
XX
XX
CC The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a genomic sequence encoding a
CC protein of the invention
XX
XX
SQ Sequence 13862 BP; 3188 A; 3560 C; 4465 G; 2649 T; 0 U; 0 Other;
Query Match 10.5%; Score 295.2; DB 4; Length 13862;
Best Local Similarity 74.6%; Pred. No. 5,9e-42;
Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;
QY 1667 AAACAGATATTAAGCTGCGAACCGTGGACACGCTCATAGTTCTAGCTGCTGGAGG 1726
DB 1508 AAACCTTAATAGCTGGCTGGGCTCTGTGGCTCATGCTGTGATCCAGACATTGGAGG 1567
QY 1727 TTGAGGAGGAGGAGTGGCTTGAACACAGCTGTCAAGGCGAGCTGGGACATACAG 1786
DB 1568 CTGAGGAGGAGGAGTCTTGAAGTCAAGGTTTGAGACCAAGCTGACCACTGGCGAA 1627
QY 1787 ATCTGTCTCTCAAAAAAAAAAAAAAAAAAGAAAGAGAGGCGCGGCGCTGGTGGC 1845
DB 1628 ACCCACTCTCTACATTAATAAATACAAAGTTAGCCAGGATGGCGGCGATGGGTGGC 1687
QY 1846 TCACGCTGTAAATCCAGCATTTGGAGAGCCGAGCGCGGATCACTGTGGTCAGA 1905
DB 1688 TCATGCTGTAAATCCAGCATTTGGAGAGCCGAGGATGGCGGATCACTGTAGGTGAGA 1747

```

```

QY 1906 GTTTGAGACCAAGCTGGCGCAATGGCAAAACCCCTGTGACTCAAAATGCAAAATTA 1965
DB 1748 GTTCAAGACCAAGCTTGGACAAATGGCAAAACCCCTGTGACTCAAAATTA 1807
QY 1966 GCCAGGCGTGTAGCAGGACCTGTAAATCCAGCTACTTGGAGAGCTGAGGACGAGAAAT 2025
DB 1808 GCCAGGCTTATGTGTAGGCGCTTAAATCCAGCTACTTGGAGAGCTGAGGACGAGAAAT 1867
QY 2026 CGCTTGAACCCAGAGAGTGGAGGTGACAGTAAGCTGAGATCGTCCGCTGCACTCCAGCC 2085
DB 1868 CGTTGAAGCCAGAGAGGAGGAGTTTGACCGAGCGAGATTTGTCACCTGCATCCAGCC 1927
QY 2086 TGGGCGACAAAGACAAAGCTTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 2145
DB 1928 TGGGCGAATTA-AGCAGAGCTGTGTCTCAGAAAAAAGAGAGAGAGAGAGAA 1986
QY 2146 AGAGACAAATATTTGGGAGAGAGATGGGAGACATTGCAAGGAAATTGG 2197
DB 1987 AGAAAAATTAGCCAGCGTGTGCAATGACCTGTAGTCTTAGCTACTTGG 2038

RESULT 15
AA07516
ID AA07516 standard; DNA; 13862 BP.
XX
XX AA07516;
AC
XX
XX 21-NOV-2001 (first entry)
DT
XX
XX Human reproductive system related antigen DNA SEQ ID NO: 10204.
DE
XX Human reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
XX Homo sapiens.
OS
XX
XX WO20015320-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001339.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225475P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.

```



```
QY 1906 GTTTGAGACCAAGCTGGCCACATGCGAATAACCCCGTCTGTACTCAAAATGCAAAATTA 1965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1748 GTTCACAGACCAAGCTGGACACATGGCAAAACCCGTGTCTATTAATAACAAAAATTA 1807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1966 GCCAGGCGTGTAGCAGGACCTGTATATCCAGCTACTTGGAGGCTGAGGCAAGAGAAT 2025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1808 GCCAGGCTTATGTAGTAGGCGCCTATATATCCAGCTACTTGGAGGCTGAGGCAAGAGAAT 1867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2026 CGCTTGAAACCCAGAGAGTGAAGGTTGAGTAAGCTGAGATCGTGCCTGTCACCTCAGCC 2085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1868 CGTTGAAAGCCAGAGAGGCGAGATTGTCACCGAGCGAGATGTGCCACTGCACCTCAGCC 1927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2086 TGGGCGACACAGAGCAAGACTTGTGTCTCAGAAAAAAGAGAGAGAGAGAGA 2145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1928 TGGGCGATTA-AGCGAGACTGTGTCTCAGAAAAAAGAGAGAGAGAGAGA 1986
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2146 AGAGAACATATTTGGAGAGAAAGATGGGGAAGCATTCAGAGAAATTTGTG 2197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1987 AGAAAAAATTAGCGAGCGTGTGTGGCATGCACTGTAGTCTTACTTGGG 2038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 7, 2004, 06:55:59  
Job time : 1065 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 06:31:31 ; Search time 217 Seconds

(without alignments)  
7206.689 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818  
Sequence: 1 agagaacatcatatga.....aaaaaaaaaaaaaaaa 2818

Scoring table: IDENTITY NUC  
Gapco 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	280	9.9	15297	4 US-09-817-180-3	Sequence 3, Appl1
2	280	9.9	15297	4 US-10-003-295-3	Sequence 3, Appl1
3	274.2	9.7	11811	3 US-09-078-294-7	Sequence 7, Appl1
4	269.8	9.6	43950	4 US-09-735-934A-3	Sequence 3, Appl1
5	269.8	9.6	43950	4 US-10-060-332-3	Sequence 3, Appl1
6	269.2	9.6	99500	4 US-09-798-096-10	Sequence 10, Appl1
7	267.4	9.5	59065	4 US-09-813-817-3	Sequence 3, Appl1
8	267.4	9.5	59065	4 US-09-978-197-3	Sequence 3, Appl1
9	265.6	9.4	9365	4 US-09-608-285A-8	Sequence 8, Appl1
10	265.6	9.4	9365	4 US-09-350-836B-8	Sequence 8, Appl1
11	265.6	9.4	9365	4 US-09-370-265-8	Sequence 8, Appl1
12	265.6	9.4	9365	4 US-09-557-800C-8	Sequence 8, Appl1
13	265.6	9.4	9365	4 US-09-570-625A-8	Sequence 8, Appl1
14	265.6	9.4	14747	4 US-09-608-285A-42	Sequence 42, Appl1
15	265.6	9.4	14747	4 US-09-557-800C-42	Sequence 42, Appl1
16	265.6	9.4	15977	4 US-09-608-285A-59	Sequence 59, Appl1
17	262.2	9.3	21234	4 US-09-810-671-3	Sequence 3, Appl1
18	262.2	9.3	21234	4 US-10-109-854-3	Sequence 3, Appl1
19	260.4	9.2	63000	4 US-09-780-172-18	Sequence 18, Appl1
20	259.2	9.2	3885	1 US-08-688-145-1	Sequence 1, Appl1
21	258.6	9.2	43950	4 US-09-735-934A-3	Sequence 3, Appl1
22	258.6	9.2	43950	4 US-10-060-332-3	Sequence 3, Appl1
23	257.6	9.1	99500	4 US-09-798-096-10	Sequence 10, Appl1
24	257.2	9.1	10684	3 US-08-618-100B-3	Sequence 3, Appl1
25	256.8	9.1	8174	1 US-07-914-281-5	Sequence 5, Appl1
26	256.8	9.1	8174	1 US-08-393-246-5	Sequence 5, Appl1
27	256.8	9.1	8174	1 US-08-525-058A-5	Sequence 5, Appl1

C 28	256.8	9.1	8174	2 US-08-696-731-5	Sequence 5, Appl1
C 29	256.8	9.1	8174	5 US-09-042-531-5	Sequence 5, Appl1
C 30	256.8	9.1	8174	5 PCT-US91-00895-3	Sequence 3, Appl1
C 31	256.8	9.1	29629	4 US-09-729-995-3	Sequence 3, Appl1
C 32	256.8	9.1	29629	4 US-10-135-689-3	Sequence 3, Appl1
C 33	254.8	9.0	118067	4 US-09-497-855A-32	Sequence 32, Appl1
C 34	254.4	9.0	98844	4 US-09-791-211-10	Sequence 10, Appl1
C 35	253.8	9.0	16053	4 US-09-801-052-3	Sequence 3, Appl1
C 36	253.8	9.0	16053	4 US-10-020-121-3	Sequence 3, Appl1
C 37	253.2	9.0	14581	4 US-08-520-373D-4	Sequence 4, Appl1
C 38	253.2	9.0	22481	4 US-08-367-841A-43	Sequence 43, Appl1
C 39	253.2	9.0	22481	5 PCT-US95-07201-43	Sequence 43, Appl1
C 40	253.2	9.0	22484	4 US-09-875-223-2	Sequence 2, Appl1
C 41	253.2	9.0	22484	4 US-09-875-114-2	Sequence 2, Appl1
C 42	252.2	8.9	7210	2 US-08-257-9635-10	Sequence 10, Appl1
C 43	252.2	8.9	7210	4 US-08-367-841A-10	Sequence 10, Appl1
C 44	252.2	8.9	7210	5 PCT-US95-07201-10	Sequence 10, Appl1
C 45	252.2	8.9	14581	4 US-08-520-373D-4	Sequence 4, Appl1

## ALIGNMENTS

RESULT 1  
US-09-817-180-3  
Sequence 3, Application US/09817180  
Patent No. 6340584  
GENERAL INFORMATION:  
APPLICANT: GAN, Weiniu et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
FILE REFERENCE: C1001183  
CURRENT APPLICATION NUMBER: US/09/817,180  
CURRENT FILING DATE: 2001-03-27  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 15297  
TYPE: DNA  
ORGANISM: Human  
US-09-817-180-3

Query Match 9.9%; Score 280; DB 4; Length 15297;  
Best Local Similarity 74.9%; Pred. No. 3e-52;  
Matches 370; Conservative 0; Mismatches 110; Indels 14; Gaps 1;

QY	1641	ATTAACTATATTAAGAGATTGTGACAAACAGATATTAAGCTGGACCGTGGACAC	1700
DB	11651	ACTCCCTTAATGCGCAACCTTCCACACAGGACGAGATTAAGATTAACCTGGCAGTGTCCAC	11710
QY	1701	GCTCATAGTCTGCTGCTGCTGGAGGTGAGAGGAGAGATGCTTGAACACAGGTTC	1760
DB	11711	GCTGTATATCCAGCACTTTGGAGGCTGAGCTGGATCACTTGAGCCAGAGTTC	11770
QY	1761	AAGGACGCTGGGCAACATTAACAAGTCTGCTCAAAAAAAAAAAAAAAAAAAGA	1820
DB	11771	AAGATCAGCTTGAGCAACAGAGAACTCATCTGTACAAAAAATACAAAAATTA	11825
QY	1821	AAGAGAGAGGCGCGGGGCTGAGCTCAGCCCTGTATCCAGCACTTTGGAGGCCGAG	1880
DB	11826	-----GACTGGGACGAGTGGCTCAGCACTGTATCCAGCACTTTGGAGGCCGAG	11876
QY	1881	CCGAGCGGATCAGCTGTGTGTGAGAGTTGAGCCAGCTGGGCAACATGGCAAAACCC	1940
DB	11877	GAGAGTGATCAGCTGTGTGTGAGAGTTGAGCCAGCAACATGGTAAACCCC	11936
QY	1941	GTCGTACTCAAAATGCAAAATTTAGCAGGCGGTGAGAGGACACTGTATCCAGCT	2000
DB	11937	ATCTTCTTAAATATCAAAATTTAGCAGGCAATGTGGACCTGTATCCAGCT	11996
QY	2001	ACTTGGAGGCTGAGGACGAGGATGCTTGAACCAAGAGGTGAGGTTGACATAGCT	2060

Db 11997 ACTTGAGAGCTGAGGTGGAGAAATTGCTTGAACCCAGAGCGGAGGCTGACAGTGAACC 12056  
QY 2061 GAGATGTCGCCGTTGACCTCCAGCCTGGGCGACAAAGAGCAAGACTTCTCTCAGAAAAA 2120  
Db 12057 GAGATTGTCCACTGCACTCCAGCCTGGGCGACAAAGAGTAAATCCATCTCAAAAAA 12116  
QY 2121 AAAAAAAAAAGAGA 2134  
Db 12117 CCAAAAAACAAAA 12130

## RESULT 2

US-10-003-295-3  
; Sequence 3, Application US/10003295  
; Patent No. 6686187  
; GENERAL INFORMATION:  
; APPLICANT: GAN, Weidhu et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEIN, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL001183DIV  
; CURRENT APPLICATION NUMBER: US/10/003,295  
; CURRENT FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 15297  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-003-295-3

Query Match 9.9%; Score 280; DB 4; Length 15297;  
Best Local Similarity 74.9%; Pred. No. 3e-52;

Matches 370; Conservative 0; Mismatches 110; Indels 14; Gaps 1;

QY 1641 ATTAACTATTTAAGAGATTGTGACAAAACAGATGATTAAGCTGGAAACCTGGCACAC 1700  
Db 11651 ACTCCCTTAATGCAACCTTCCACAGAGAGATGAATTAAGTAACTGGCCAGTTGCTCAC 11710  
QY 1701 GCTCATAGTTCTAGCTGCTTGGAGGTTGAGAGGAGAGATGGCTTGAACACAGGTTTC 1760  
Db 11711 GCTGTGATCTCCAGCACTTGGAGGCTGAGCTGGTGAATCACTTGAAGCCAGAGTTTC 11770  
QY 1761 AAGCCAGCTTGCGCAACATTAACAAGATCTGTCTCTCAAAAAAATAAAAAAAGA 1820  
Db 11771 AAGATCAGCTTGCAACACAGTGAATCTCTTAACAAAAAATAAAAAAATA----- 11825  
QY 1821 AAGAGAGAGGCGCGGCGTGTGCTGACGCTGTAAATCCAGCACTTGGAGGCGGAG 1880  
Db 11826 -----GACTGGGACCGGTGGCTCACACCTGTAACTCCAGCACTTGGAGGCGGAG 11876  
QY 1881 CCGGCGCGATCACCTGTGTGACAGAGATTGAGCCAGCCTGGCCACATGGCAAAACCCC 1940  
Db 11877 GCAAGGTGATCACCTGTGTGACAGAGATTGAGCCAGCCTGGCCACATGGTGAACCCC 11936  
QY 1941 GTCGTACTCAAAATGCAAAATTTAGCAGGCGTGTAGAGGCACTGTAAATCCAGCT 2000  
Db 11937 ATCTCTACTAAATTAACAAATTTAGCAGGCAATGTGTGACGCTGTAAATCCAGCT 11996  
QY 2001 ACTTGGAGGCTGAGGAGAGAGATGCTTGAACCCAGAGGTGAGGTTGCAATAGCT 2060  
Db 11997 ACTTGGAGGCTGAGGAGAGATGCTTGAACCCAGAGGTGAGGTTGCAATAGCT 12056  
QY 2061 GAGATGTCGCCGTTGACCTCCAGCCTGGGCGACAAAGAGCAAGCTCTCTCAGAAAAA 2120  
Db 12057 GAGATTGTCCACTGCACTCCAGCCTGGGCGACAAAGAGTAAATCCATCTCAAAAAA 12116  
QY 2121 AAAAAAAAAAGAGA 2134  
Db 12117 CCAAAAAACAAAA 12130

RESULT 3

## US-09-078-294-7/c

; Sequence 7, Application US/09078294  
; Patent No. 6265211  
; GENERAL INFORMATION:  
; APPLICANT: Choo, Kong-Hong Andy  
; APPLICANT: Du Sart, Desiree  
; APPLICANT: Cancilla, Michael R.  
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
; FILE REFERENCE: Davies Col  
; CURRENT APPLICATION NUMBER: US/09/078, 294  
; CURRENT FILING DATE: 1998-05-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 11811  
; TYPE: DNA  
; ORGANISM: BAC-F2 contig 3  
US-09-078-294-7

Query Match 9.7%; Score 274.2; DB 3; Length 11811;  
Best Local Similarity 79.7%; Pred. No. 5.2e-51;

Matches 350; Conservative 0; Mismatches 83; Indels 6; Gaps 2;

QY 1692 GTGGCAAGCTCATAGTTGTTAGCTGCTGGAGGTTGAGAGGAGAGATGGCTTGAACA 1751  
Db 6663 GTGGCTATGCTGTGTAATCCAGCACTTGGAGGCGCGAGGTGGAGAAATGACTTGAAGCT 6604  
QY 1752 CAGGTTTCAAGGCCAGCCTGGGCAACATTAACAAGATCTGTCTCAAAAAAATAAAAAA 1811  
Db 6603 CAGAGTTCAAGACAGCCTGGGCAACAGTGAAGCTCATCTCAAAAAATACATTTA 6544  
QY 1812 AAAAAAAAAAGAGAGAGGCGCGGCGTGTGCTCAAGCTGTAAATCCAGCACTTTGG 1871  
Db 6543 AAGATTAG-----CTGGGCGCGAGGTGTGTGCGCACCTGTAAATCCAGCCTTTGG 6489  
QY 1872 GAGGCCAGCGCGGAGATCACTGTGTGAGAGTTGAGACCACTGGCCCAACATGG 1931  
Db 6488 GAGGTCAAGGTGTGTGATCACTGAGGTGAGAGTTGAGGCCAGCCTGGCCCAACATGG 6429  
QY 1932 CAAGACCCCTGTGTAATCAAAATGC-AAAATTAACCAAGCCTGTGTAGCAGCACTGT 1990  
Db 6428 TGAAGCCCACTCTCTACTAATAAATAAATAAATTAACCAAGCCTGTGTAGCAGCACTGT 6369  
QY 1991 AATCCAGCTACTTGGAGGCTGAGGAGAGAGATGCTTGAACCCAGAGAGTGAAGTT 2050  
Db 6368 AATCCAGCTACTGAGAGGCTGAGGAGAGAGATGCTTGAACCCAGAGAGAGTGAAGTT 6309  
QY 2051 GCAATAGCTGAGATGCTGCGCTGTGCACTCCAGCCTGGGCGACAAAGAGCAAGCTGTTC 2110  
Db 6308 GCACTGAGCCAAAGTCAATGCACTTGAAGTCACTCCAGCCTGGGCGACAAAGAGCAAGCTGTCA 6249  
QY 2111 TCAGAAAAAATAAAAAA 2129  
Db 6248 AAAAAAAAAAAAAA 6230

## RESULT 4

US-09-735-934A-3/c  
; Sequence 3, Application US/09735934A  
; Patent No. 6372468  
; GENERAL INFORMATION:  
; APPLICANT: Li, Jia-Yin et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL000851  
; CURRENT APPLICATION NUMBER: US/09/735, 934A  
; CURRENT FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 43950  
; TYPE: DNA

RESULT 4



ORGANISM: Homo sapiens  
US-09-735-934A-3

US-09-735-934A-3

Query Match	9.6%	Score 269.8;	DB 4;	Length 43950;
Best Local Similarity	77.9%	Pred. No. 8.1e-50;		
Matches 339; Conservative	0;	Mismatches 92;	Indels 4;	Gaps 1;

Qy	1685	GCGAACCCTGGCACAACCCCTCATAGTTCTAGCTGCTTGGAGAGTTGAGAGGGAGATGGC	1744
Db	17761	GGGCAAGTGGCACAACACTGTATCCGABAATTGGAGGCTGAGGTGGAGAGATTGC	17702
Qy	1745	TTGAACAACAGTGTTCACAGGCCAGCTGGGCAACATPACAGATCCTGTCTCTAAAAA	1804
Db	17701	TTGGGCCAGAGTTCAAGACCAACCCCTGGGCAACATGACAAAGCCCCCTATCTTTAAAAA	17642
Qy	1805	AAAAAAAAAAAAAAAAGA----AAGAGAGAGGCGCGGGGCTGTGTGGTCAAGCCCTGTAATGC	1860
Db	17841	AAAAAAAAAAAAAAAAAAGAGAGATATAGGTCAAGGCGCGGTGCTCATAGCTCTGTATCC	17582
Qy	1861	CAGCACTTTGGGAGGCGGAGCGCGGCGGATCACTGTGTCAAGAATTGAGACCAAGCT	1920
Db	17581	CAGCACTTTGGGAGGCGGAGCGGAGGTGATTACTGTAGTTGGAGATTCAAGCAAGCT	17522
Qy	1921	GGCCACAATGGCAAAACCCCTGTCTGTACTAAAAATGCAAAATTAAGCAGGCGGTGTAC	1980
Db	17521	GACCAACAATGGCAAAACCCCTGTCTCTCTAAAAATCAAAAAATGATGCTGGGCTGGGC	17462
Qy	1981	AGGCACCTGTATCCAGCTACTCTGGGAGGCTGAGCAGAGAAATGCTTGAACCCAGGA	2040
Db	17461	GGGTGCTCTGTAGTCCCAAGCTACTTTGGAGGCTGAGCAGAGAAATGCTTGAACCCAGGA	17402
Qy	2041	GGTGAGAGTTGCAGTAAGCTGAGATCGTGCCTGTGCACTTCAGCCTGGGCGACAAAGCA	2100
Db	17401	GACAGAGTTGCATGAATGAATTCATCACTGCACTCAACCTGGGCAACAAAGCG	17342
Qy	2101	AGACTGTCTCAGA 2115	
Db	17341	AAACGCCGTCTAAAA 17327	

RESULT 5  
US-10-060-332-3/c

```

? Sequence 3, Application US/10060332
? Patent No. 6528294
? GENERAL INFORMATION:
? APPLICANT: LI, JIAYIN et al
? TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
? TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
? TITLE OF INVENTION: THEREOF
? FILE REFERENCE: CLO00851DIV
? CURRENT APPLICATION NUMBER: US/10/060,332
? CURRENT FILING DATE: 2002-02-01
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 3
? LENGTH: 43950
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-10-060-332-3

```

Query Match	9.6%	Score 269.8	DB 4	Length 43950
Best Local Similarity	77.9%	Pred. No. 8.1e-50		
Matches 339	Conservative 0	Mismatches 92	Indels 4	Gaps 1
Qy	1685	GCGAACCCTGGCAACGCTCATAGTTTCTTACCTGTTGGAGCTTGGAGAGGAGCATGCG	1744	
Db	17761	GGGCAACATGGCAACACCTGTATCCCAAACTTTGGAGGCTGAAGTGGAGGATTCG	17702	
Qy	1745	TTGAACAAGGTGTTCAAGGCACGCTTGGCAACATACAGATCCTGTCTCAAAAA	1804	
Db	17701	TTGAGCGCAAGATTTCAAGACCACTCTGGGCAACATAGCAAGACCCCTTATCTTAAAAA	17642	
Qy	1805	AAAAAAAAAAAAAGA---AGAGAGAGAGGCGCGCGTGTGGCTTCAACGCTGTATCC	1860	

Db	17641	AAAAAAAAAAAAAAAAAGAGAGAAATATAGGTAGAGCGCGGTGGCTCATGCTCTGTATCC	175823
Qy	1861	CAGCACTTTGGAGGCGGAGCCGGGCGGATCACCTTGTGTGAGAGTTTGAACCGCTT	1920
Db	17581	CAGCACTTTGGAGGCGCGAGGCGAGGTGATTTACTCTAGGTTGGGAGTTCAAGCCGAGCTT	175222
Qy	1521	GGCCCAATGGCAAAACCCCGTCTGTATCTCAAAATGCAAAATTTGACGAGGGGTGTAC	1980
Db	17521	GACCAATCATGGCAAAACCTGTCTCTATCAAAATCAAAAGTAAAGTGGGCTTGGGCT	17462
Qy	1981	AGGCACTGTATATCCAGCTACTTTGGGAGGCTGAGGCAAGAGATCGTTGAACCCAGGA	2040
Db	17461	GGGTGCTGTATGTCCAGCTACTTTGGGAGGCTGAGGCAAGAGATCGTTGAACCCAGGA	17402
Qy	2041	GGTGAGAGTTGCAGTAAGCTGAATGTGTGCCCTTGCATCTCCAGCTGGGCGCAAGAGGA	2100
Db	17401	GACAGAGTTGCAGTAAGCTGAATTTGATCACTGCACTCCAGCTGGGCGCAAGAGGCG	17342
Qy	2101	AGACTGTGTCTAGA 2115	
Db	17341	AAAGCGGCTTAAAA 17327	

```

RESULT 6
US-03-798-096-10/c
/ Sequence 10 Application US/03798096
/ Patent No. 6399378
/ GENERAL INFORMATION:
/ APPLICANT: Donna T. Ward
/ APPLICANT: Andrew T. Walt
/ TITLE OF INVENTION: ANTISENSE MODULATION OF RECOL2 EXPRESSION
/ FILE REFERENCE: RTS-0207
/ CURRENT APPLICATION NUMBER: US/03/798,096
/ CURRENT FILING DATE: 2001-03-01
/ NUMBER OF SEQ ID NOS: 89
/ SEQ ID NO 10
/ LENGTH: 99500
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
US-03-798-096-10

```

```

Query Match          9.68; Score 269.2; DB 4; Length 99500;
Best Local Similarity 73.7%; Pred. No. 1.5e-49;
Matches 386; Conservative 0; Mismatches 103; Indels 35; Gaps 2;

QY      1661  GTGACAAACAGATGATTAAGCTGGCAACCGTGGCACACGCTCATAGTCTTAGCTCTT 1720
DB      43878  GGGAAAAAAGCTTGGCTTAGGACATGGTGGCTCATGACTGTAAATCCAGCCCTTT 43819

QY      1721  -GGAGGTTAGAGAGGAGGATGGCTTTGAACAGAGTGTTCAAAGGCAAGCTGGGCAACA 1779
DB      43818  GGGAGGCCAAGGAGGAGGATGTCTTAAGTCCAGAGTTCAAGACACACCTGGGCAACA 43758

QY      1780  TAAACAATCTGTCTCTCAAAAAAAAAAAAAAAAAAAAAAGAGAG----- 1828
DB      43758  CAGTGAACCTGTCACTAAAAAAAAAAAAAAAAAGCTTAAGTCTTGATTTCTTTTCT 43699

QY      1829  -----GGGCGGGCGGTGGTGGCTCAGGCTGTAAATCCAGCA 1865
DB      43688  TTTTTTTTTTTTTTAAAGTGGGGCTGGGCAAGGTGGCTCAGGCTGTAAATCCAAAC 43633

QY      1866  CTTTGGAGGCGCAGCGCGCGGATCACTGTGTGAGAGGTTTGAACAGCCTTGCCA 1925
DB      43638  CTTTGGAGGCGCAAGAGGTTAATCACTGTAGGTGAGAGGTTCCAGACCAAGCTGGCCA 43579

QY      1926  ACATGGCAAAACCCCGCTGTGTACTCAAAATGCAAAATTAGCCAGGCGGTGTAGCAGCA 1985
DB      43578  ATGTGTGAACCCCGCTTTACTAATTAATATAAAATTAGCGGCGCATGGTGTGCTGTG 43518

QY      1986  CTTGTATCCCATGTAATTGGAGGCTGAGGCAAGAGCAATCGTTGAACCCAGAGGTGG 2045
DB      43518  CTTGTATCCCATGTAATTGGAGGCTGAGGCAAGAGCAATCGTTGAACCCAGAGGTGG 2045

```

Db 43518 TCTGTATCCAGCTACTGCGGAGGCTGAGGAGGAGATCGCTTGAACCTGGAGGCTGG 43459  
QY 2046 AGGTTCAGTAACTGAGATCGTGCCTTGCCTCCAGCTGGCGGACAAAGAGAGT 2105  
Db 43458 AGGTTCAGTAACTGAGATCGTGCCTTGCCTCCAGCTGGCGGACAAAGAGT 43399  
QY 2106 CTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAG 2149  
Db 43398 TTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAG 43355

## RESULT 7

US-09-813-817-3/c  
Sequence 3, Application US/09813817  
Patent No. 6340583  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEIN, NUCLEIC  
ACID MOLECULES ENCODING HUMAN KINASE PROTEIN, AND USES  
FILE REFERENCE: C1001178  
CURRENT APPLICATION NUMBER: US/09/813,817  
CURRENT FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 59065  
TYPE: DNA  
ORGANISM: Human  
US-09-813-817-3

Query Match 9.5%; Score 267.4; DB 4; Length 59065;  
Best Local Similarity 78.2%; Pred. No. 3.1e-49;  
Matches 351; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1680 AAGCTGGAACCGTGGACACGCTCAATGTTCTAGCTGCTGGAGGTTGAGAGGAGG 1739  
Db 23369 AGCCAGGCAATAGTGTGCTCATGCTTAATCCAGC-ATTGAGAGGCGCAAGAGGAGG 23311  
QY 1740 ATGGCTTGAACACAGGTGTTCAAGGCGCAGCTGGGCAATACAAAGATCTGTCTCA 1799  
Db 23310 ACTGCTTGAAGCCAGAAAGTTTGAGACAGCTGGGTAACAAACCAAGACCTTATCTTAC 23251  
QY 1800 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGGCGTGGCTCAGCCTGTATTC 1859  
Db 23550 TAAAAAATAGTCAATCA-----GGCCAGTCAATGTGGCTCAGCCTGTATTC 23202  
QY 1860 CAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTGTCAGAGTTTGAAGCCAGC 1919  
Db 23201 CAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTGTCAGAGTTTGAAGTTGAGC 23142  
QY 1920 TGGCCAAATGAGCAAAACCCGCTGTGTACTCAAAATTCGCAAGGCGCTGTAG 1979  
Db 23141 TGACCAACATGAGAAACCTGTGTCTCTAATAAAATACAAAAATTGCGCAGGCGTGTGG 23082  
QY 1980 CAGGCACTGTATTCAGCTACTTGGAGGCTGAGGAGGAGATCGCTTGAACCCAGG 2039  
Db 23081 CAGTGTCTGTATTCAGCTACTTGAAGGCTGAGAGCAGAGAAATCGCTTGAACCCAGG 23022  
QY 2040 AGGTGAGGTTGAGTAACTGAGATCGTGCCTTGCCTCCAGCTGGGCGACAGAGC 2099  
Db 23021 AGGAGAGGTTGGGAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 22962  
QY 2100 AAGACTGTGTCTCAGAAAAAAGAAAAA 2128  
Db 22961 GAAACTGTGTCTCAAAAAAGAAAAA 22933

RESULT 8  
US-09-978-197-3/c  
Sequence 3, Application US/09978197  
Patent No. 6403353  
GENERAL INFORMATION:

APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEIN, NUCLEIC  
ACID MOLECULES ENCODING HUMAN KINASE PROTEIN, AND USES  
FILE REFERENCE: C1001178D10  
CURRENT APPLICATION NUMBER: US/09/978,197  
CURRENT FILING DATE: 2001-10-17  
PRIOR APPLICATION NUMBER: 09/813,817  
PRIOR FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 59065  
TYPE: DNA  
ORGANISM: Human  
US-09-978-197-3

Query Match 9.5%; Score 267.4; DB 4; Length 59065;  
Best Local Similarity 78.2%; Pred. No. 3.1e-49;  
Matches 351; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1680 AAGCTGGAACCGTGGACACGCTCAATGTTCTAGCTGCTGGAGGTTGAGAGGAGG 1739  
Db 23369 AGCCAGGCAATAGTGTGCTCATGCTTAATCCAGC-ATTGAGAGGCGCAAGAGGAGG 23311  
QY 1740 ATGGCTTGAACACAGGTGTTCAAGGCGCAGCTGGGCAATACAAAGATCTGTCTCA 1799  
Db 23310 ACTGCTTGAAGCCAGAAAGTTTGAGACAGCTGGGTAACAAACCAAGACCTTATCTTAC 23251  
QY 1800 AAAAAAAAAAAAAAAAAAGAGAGAGGCGCGGCGTGGCTCAGCCTGTATTC 1859  
Db 23550 TAAAAAATAGTCAATCA-----GGCCAGTCAATGTGGCTCAGCCTGTATTC 23202  
QY 1860 CAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTGTCAGAGTTTGAAGCCAGC 1919  
Db 23201 CAGCACTTTGGAGGCGGAGCGGCGGATCACTGTGTGTCAGAGTTTGAAGTTGAGC 23142  
QY 1920 TGGCCAAATGAGCAAAACCCGCTGTGTACTCAAAATTCGCAAGGCGCTGTAG 1979  
Db 23141 TGACCAACATGAGAAACCTGTGTCTCTAATAAAATTCGCAAGGCGCTGTAG 23082  
QY 1800 CAGGCACTGTATTCAGCTACTTGGAGGCTGAGGAGGAGATCGCTTGAACCCAGG 2039  
Db 23081 CAGTGTCTGTATTCAGCTACTTGAAGGCTGAGAGCAGAGAAATCGCTTGAACCCAGG 23022  
QY 2040 AGGTGAGGTTGAGTAACTGAGATCGTGCCTTGCCTCCAGCTGGGCGACAGAGC 2099  
Db 23021 AGGAGAGGTTGGGAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 22962  
QY 2100 AAGACTGTGTCTCAGAAAAAAGAAAAA 2128  
Db 22961 GAAACTGTGTCTCAAAAAAGAAAAA 22933

## RESULT 9

US-09-608-285A-8/c  
Sequence 8, Application US/09608285A  
Patent No. 633013  
GENERAL INFORMATION:  
APPLICANT: Ford, John  
APPLICANT: Mulero, Julio  
APPLICANT: Yeung, George  
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
POLYPEPTIDES  
FILE REFERENCE: 28110/36570  
CURRENT APPLICATION NUMBER: US/09/608,285A  
CURRENT FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: 09/583,231  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 09/557,800  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/481,238  
PRIOR FILING DATE: 2000-01-11

```

; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-608-285A-8

Query Match          9.4%; Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 3,8e-49;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1765 CCGAGCTGGGCAACATACAGATCTCTCTCAAAAAAAAAAAAAAAAAAG- 1823
DB 7802 CCGAGATGGGTGACAGAGCAAGACCTCTCTCAAAAAAAAAAAAAAAAAAG 7743
QY 1824 -----AGAGAGGCGCGGCGGTGCTCAGCGCTGTATCCAGACCTTTGGAGG 1875
DB 7742 AAMCGAMCAAAAAGGCTRGAGTGTGCTCAATCTCTATCTTACCACTTTGGAGG 7683
QY 1876 CCGAGCCGGGCGGATCAGCTGTGTGTCAGAGAGTTTGAACAAGCTTGCCAAATGGCATA 1935
DB 7682 CYGAGGCGGTGATCACTGAGTCAAGATGTTTGAAGTTCAGAGAGGCTTGCAATGGTGA 7623
QY 1936 ACCCGCTCTGTACTCAAAATGC--AAAAATTAGCAGGCGGTGTAGCAGGCACTGTAT 1993
DB 7622 ACCCGCTCTGTACTCAAAATGC--AAAAATTAGCAGGCGGTGTAGCAGGCACTGTAT 7563
QY 1994 CCGAGTACTTGGAGGCGGTGAGGAGGAATCGTTAAACCGAGAGTGTGAGTTGCA 2053
DB 7562 CCGAGTACTTGGAGGCGGTGAGGAGGAATCGTTAAACCGAGAGGAGTGTGCA 7503
QY 2054 GTAAGCTGAGATGTCGCTGTCACTCCAGCTGGGCAACAAGAGCAAGCTGTGTCA 2113
DB 7502 GTGAGCTGAGATGTCGCTGTCACTCCAGCTGGGCAACAAGAGCAAGCTGTGTCA 7443
QY 2114 GAAAAAAAAAAAAAAAAA 2129
DB 7442 AAAAAATATATATAA 7427

RESULT 10
US-09-350-836B-8/c
; Sequence 8, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
```

```

; APPLICANT: Ferd. John
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
; NAME/KEY: exon
; LOCATION: (1281)..(1580)
; NAME/KEY: exon
; LOCATION: (1820)..(1855)
; NAME/KEY: exon
; LOCATION: (2467)..(2555)
; NAME/KEY: exon
; LOCATION: (2863)..(2942)
; NAME/KEY: exon
; LOCATION: (3889)..(3950)
; NAME/KEY: exon
; LOCATION: (4894)..(4995)
; NAME/KEY: exon
; LOCATION: (5847)..(5987)
; NAME/KEY: exon
; LOCATION: (6966)..(7138)
; NAME/KEY: exon
; LOCATION: (8556)..(9365)
; NAME/KEY: misc.feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc.feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc.feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
; NAME/KEY: misc.feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-350-836B-8

Query Match          9.4%; Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 3,8e-49;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1765 CCGAGCTGGGCAACATACAGATCTCTCTCAAAAAAAAAAAAAAAAAAG- 1823
DB 7802 CCGAGATGGGTGACAGAGCAAGACCTCTCTCAAAAAAAAAAAAAAAAAAG 7743
QY 1824 -----AGAGAGGCGCGGCGGTGCTCAGCGCTGTATCCAGACCTTTGGAGG 1875
DB 7742 AAMCGAMCAAAAAGGCTRGAGTGTGCTCAATCTCTATCTTACCACTTTGGAGG 7683
QY 1876 CCGAGCCGGGCGGATCAGCTGTGTGTCAGAGAGTTTGAACAAGCTTGCCAAATGGCATA 1935
DB 7682 CYGAGGCGGTGATCACTGAGTCAAGATGTTTGAAGTTCAGAGAGGCTTGCAATGGTGA 7623
QY 1936 ACCCGCTCTGTACTCAAAATGC--AAAAATTAGCAGGCGGTGTAGCAGGCACTGTAT 1993
```



```

; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc_feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-557-800C-8

Query Match          9.4%; Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 3.8e-49;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1765 CCAGCCTGGGCAACATACAGATCCTGCTCTCAAAAAAAAAAAAAAAAAAGAAAG-1823
DB 7802 CCAGCATGGGTGACAGAGCAAGACCTGTCTCAAAAAAAAAAAAAAAAAAG 7743
QY 1824 -----AGAGAGGCGCGGCGTGTGCTCAAGCTGTAATCCAGACCTTGGAGG 1875
DB 7742 AAMGAMCAAAAAAAAAAGGCTRGAGGTGTGCTCACTCTTAATCTTAAGCACTTTGGAGG 7683
QY 1876 CCGAGCCGCGGCGGATCACTGTGTGTCAGAGTTGAGACCAAGCCTGCGCAACATGCGCAA 1935
DB 7682 CYGAGGCGGAGTGAATCACTGAGRTCAAGATTYRAGAMVAGCTGCGCAACATGCTGA 7623
QY 1936 ACCCGCTGTGTCTCAAAATGC--AAAAATTAGCCAGGCTGTGACAGGACCTGTAT 1993
DB 7622 ACCCGCTGTGTCTCAAAATGC--AAAAATTAGCCAGGCTGTGCGCGCTGTAT 7563
QY 1994 CCCAGCTACTTGGAGGCTGAGGAGAGATGCTTGAACCCAGAGAGGAGTGTGA 2053
DB 7562 CCCAGCTACTTGGAGGCTGAGGAGAGATGCTTGAACCCAGAGAGGAGTGTGA 7503
QY 2054 GTAAGCTGAGATGTCGCTGTGACCTCAAGCTGCGGCAAGAGCAAGACTGTCTCA 2113
DB 7502 GTGAGCTGAGATGTCGCTGTGACCTCAAGCTGCGGCAAGAGCAAGACTGTCTCA 7443
QY 2114 GAAAAAAAAAAAAAAAA 2129
DB 7442 AAAAAAAAAATAATAA 7427

RESULT 13
US-09-370-625A-8/c
; Sequence 8, Application US/09370625A
; Patent No. 6600032
; GENERAL INFORMATION:
; APPLICANT: Ford, John
```

```

; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/09/370, 625A
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350, 836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273, 447
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1) .. (288)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1281) .. (1580)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1820) .. (1855)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2467) .. (2555)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (2863) .. (2942)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (3889) .. (3950)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (4894) .. (4995)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (5847) .. (5987)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (6966) .. (7138)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (8556) .. (9365)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3409)
; OTHER INFORMATION: n = a or g or t or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9214)
; OTHER INFORMATION: n = a or g or t or c
; NAME/KEY: misc_feature
; LOCATION: (9303)
; OTHER INFORMATION: n = a or g or t or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9311)
; OTHER INFORMATION: n = a or g or t or c
US-09-370-625A-8

Query Match          9.4%; Score 265.6; DB 4; Length 9365;
Best Local Similarity 79.0%; Pred. No. 3.8e-49;
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1765 CCAGCCTGGGCAACATACAGATCCTGCTCTCAAAAAAAAAAAAAAAAAAGAAAG-1823
DB 7802 CCAGCATGGGTGACAGAGCAAGACCTGTCTCAAAAAAAAAAAAAAAAAAG 7743
```

QY 1824 -----AGAGGCGCGCGCTGTGCTCAAGCTGTAAATCCAGACTTTGGAGG 1875  
DB 7742 AAMGAMCAAAAAAGGCTGAGGTGTGCTCAATCTTAATTCYAGACTTTGGAGG 7683  
QY 1876 CCGAGCGGCGGCGGATCACTGTGTGAGAGTTGAGACCAAGCTGGCAATGAGCAA 1935  
DB 7682 CYAGAGTGGGTGATCACTGTGATCAAGATTCAGAGTTTAAAGCTGGCCAAATGTTGA 7623  
QY 1936 ACCCGCTGTCTACTCAAAATGC-AAAAATTAGCCAGGCGGTGTAGCAAGCACTGTAA 1993  
DB 7622 ACCCGCTGTCTACTCAAAATGC-AAAAATTAGCCAGGCGGTGTAGCAAGCACTGTAA 7563  
QY 1994 CCGAGCTACTTGGAGGCTGAGGCAAGATTCCTTGAACCCAGAGGTGAGGTTGA 2053  
DB 7562 CCGAGCTACTTGGAGGCTGAGGCAAGATTCCTTGAACCCAGAGGTGAGGTTGA 7503  
QY 2054 GTAGCTGAGATGTCGCTGTGCACTCCAGCTGGGCGCAAGAGCAAGCACTGTCTCA 2113  
DB 7502 GTAGCTGAGATGTCGCTGTGCACTCCAGCTGGGCGCAAGAGCAAGCACTGTCTCA 7443  
QY 2114 GAAAAAAAAAAAAA 2129  
DB 7442 AAAAAAAAAAAAAA 7427

RESULT 14  
US-09-608-285A-42/c  
Sequence 42, Application US/09608285A  
Patent No. 6335013  
GENERAL INFORMATION:  
APPLICANT: Ford, John  
APPLICANT: Mulero, Julio  
APPLICANT: Yeung, George  
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
TITLE OF INVENTION: POLYPEPTIDES  
FILE REFERENCE: 28110/36570  
CURRENT APPLICATION NUMBER: US/09/608,285A  
CURRENT FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: 09/583,231  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 09/557,800  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/481,238  
PRIOR FILING DATE: 2000-01-11  
PRIOR APPLICATION NUMBER: 09/370,265  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: PCT/US99/16180  
PRIOR FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: 09/350,836  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: 09/273,447  
PRIOR FILING DATE: 1999-03-19  
PRIOR APPLICATION NUMBER: 09/244,444  
PRIOR FILING DATE: 1999-02-04  
PRIOR APPLICATION NUMBER: 09/122,449  
PRIOR FILING DATE: 1998-07-24  
PRIOR APPLICATION NUMBER: 09/118,205  
PRIOR FILING DATE: 1998-07-16  
NUMBER OF SEQ ID NOS: 60  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 42  
LENGTH: 14747  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (13641)  
OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine  
US-09-608-285A-42

Query Match 9.4%; Score 265.6; DB 4; Length 14747;  
Best Local Similarity 79.0%; Pred. No. 4.5e-49;  
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;

QY 1765 CCAAGCTGGGCAACATACAGATCTGTCTCAAAAAAAAAAAAAAAAAAGAAAG- 1823  
DB 11162 CCAAGATGGGTGACAGAGCAAGACCTGTCTCAAAAAAAAAAAAAAAAAAG 11103  
QY 1824 -----AAGAGGCGCGGCGGTGTGCTCAAGCTGTAAATCCAGACTTTGGAGG 1875  
DB 1102 AAMGAMCAAAAAAGGCTGAGGTGTGCTCAATCTTAATTCYAGACTTTGGAGG 11043  
QY 1876 CCGAGCGGCGGCGGATCACTGTGTGAGAGTTGAGACCAAGCTGGCAATGAGCAA 1935  
DB 11042 CYAGAGTGGGTGATCACTGTGATCAAGATTCAGAGTTTAAAGCTGGCCAAATGTTGA 10983  
QY 1936 ACCCGCTGTCTACTCAAAATGC-AAAAATTAGCCAGGCGGTGTAGCAAGCACTGTAA 1993  
DB 10982 ACCCGCTGTCTACTCAAAATGC-AAAAATTAGCCAGGCGGTGTAGCAAGCACTGTAA 10923  
QY 1994 CCGAGCTACTTGGAGGCTGAGGCAAGATTCCTTGAACCCAGAGGTGAGGTTGA 2053  
DB 10922 CCGAGCTACTTGGAGGCTGAGGCAAGATTCCTTGAACCCAGAGGTGAGGTTGA 10863  
QY 2054 GTAGCTGAGATGTCGCTGTGCACTCCAGCTGGGCGCAAGAGCAAGCACTGTCTCA 2113  
DB 10862 GTAGCTGAGATGTCGCTGTGCACTCCAGCTGGGCGCAAGAGCAAGCACTGTCTCA 10803  
QY 2114 GAAAAAAAAAAAAA 2129  
DB 10802 AAAAAAAAAAAAAA 10787

RESULT 15  
US-09-557-800C-42/c  
Sequence 42, Application US/09557800C  
Patent No. 6476211  
GENERAL INFORMATION:  
APPLICANT: Ford, John  
APPLICANT: Mulero, Julio  
APPLICANT: Yeung, George  
TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 28110/36457  
CURRENT APPLICATION NUMBER: US/09/557,800C  
CURRENT FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/481,238  
PRIOR FILING DATE: 2000-01-11  
PRIOR APPLICATION NUMBER: 09/370,265  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: PCT/US99/16180  
PRIOR FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: 09/350836  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: 09/273447  
PRIOR FILING DATE: 1999-03-19  
PRIOR APPLICATION NUMBER: 09/122449  
PRIOR FILING DATE: 1998-07-24  
PRIOR APPLICATION NUMBER: 09/244444  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 09/118,205  
PRIOR FILING DATE: 1998-07-16  
NUMBER OF SEQ ID NOS: 56  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 42  
LENGTH: 14747  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (13641)  
OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine  
US-09-557-800C-42

Query Match 9.4%; Score 265.6; DB 4; Length 14747;  
Best Local Similarity 79.0%; Pred. No. 4.5e-49;  
Matches 297; Conservative 34; Mismatches 34; Indels 11; Gaps 2;





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 09:57:52 ; Search time 964 Seconds

(without alignments)  
10691.539 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818

Sequence: 1 agagaacacatcaatctga.....aaaaaaaaaaaaaaaaaa 2818

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2818	100.0	2818	9 US-09-966-880A-7	Sequence 7, Appl1
2	2174.6	77.2	6564	9 US-09-966-880A-10	Sequence 10, Appl1
3	2174.6	77.2	11204	9 US-09-966-880A-35	Sequence 35, Appl1
4	2172	77.2	2172	9 US-09-966-880A-15	Sequence 15, Appl1
5	603.4	21.4	2440	9 US-09-966-880A-1	Sequence 1, Appl1
6	295.2	10.5	13862	10 US-09-764-891-5477	Sequence 5477, Ap
7	295.2	10.5	13862	10 US-09-764-891-10204	Sequence 10204, A
8	295.2	10.5	13862	14 US-10-205-428-1003	Sequence 1003, Ap
9	292.6	10.4	7809	10 US-09-764-891-6094	Sequence 6094, Ap
10	284.4	10.1	32349	10 US-09-764-891-7364	Sequence 7364, Ap
11	281.4	10.0	65608	9 US-09-962-436-232	Sequence 232, App
12	281.4	10.0	65608	9 US-09-962-832-119	Sequence 119, App
13	281.4	10.0	65608	9 US-09-954-531-180	Sequence 180, App
14	280.2	9.9	11821	9 US-09-764-877-2857	Sequence 2857, Ap
15	280.2	9.9	11821	15 US-10-242-515-2857	Sequence 2857, Ap

16	280	9.9	15297	13 US-10-003-235-3	Sequence 3, Appl1
17	278.2	9.9	32204	10 US-09-764-891-8537	Sequence 8537, Ap
18	277.4	9.8	30000	14 US-10-262-552-33	Sequence 33, Appl1
19	276.6	9.8	37940	15 US-10-348-073A-12	Sequence 12, Appl1
20	276.6	9.8	38000	15 US-10-175-627-11	Sequence 11, Appl1
21	275.4	9.8	3965	14 US-09-764-855-227	Sequence 227, App
22	275.4	9.8	3965	14 US-10-072-349-227	Sequence 227, App
23	275	9.8	6482	14 US-10-180-410-6	Sequence 6, Appl1
24	275	9.8	20050	14 US-10-180-410-5	Sequence 5, Appl1
25	274.8	9.8	15362	9 US-09-764-877-2856	Sequence 2856, Ap
26	274.8	9.8	15362	15 US-10-242-515-2856	Sequence 2856, Ap
27	274.2	9.7	11811	10 US-09-728-552-7	Sequence 7, Appl1
28	274.2	9.7	176001	16 US-10-210-556-27	Sequence 27, Appl1
29	274.2	9.7	186739	16 US-10-210-556-19	Sequence 19, Appl1
30	274	9.7	4963	9 US-09-764-877-2903	Sequence 2903, Ap
31	274	9.7	4963	15 US-10-242-515-2903	Sequence 2903, Ap
32	273.6	9.7	174424	10 US-09-967-768A-134	Sequence 314, App
33	273.6	9.7	174424	10 US-09-960-706-969	Sequence 969, App
34	273	9.7	2401	15 US-10-027-632-102635	Sequence 102635, Ap
35	273	9.7	26657	9 US-09-810-673A-3	Sequence 3, Appl1
36	272	9.7	17491	14 US-10-395-242-3	Sequence 3, Appl1
37	272	9.7	17491	14 US-10-017-161-1995	Sequence 1995, Ap
38	271	9.6	271	9 US-09-966-880A-13	Sequence 13, Appl1
39	270.8	9.6	670	15 US-10-027-632-201375	Sequence 201375, Ap
40	270.8	9.6	1061	14 US-10-277-161-66	Sequence 66, Appl1
41	270.8	9.6	107820	10 US-09-792-616-1	Sequence 1, Appl1
42	270.2	9.6	17286	9 US-09-764-877-3234	Sequence 3234, Ap
43	270.2	9.6	17286	15 US-10-242-515-3334	Sequence 3234, Ap
44	269.8	9.6	43950	13 US-10-060-332-3	Sequence 3, Appl1
45	269.8	9.6	43950	14 US-10-339-657-3	Sequence 3, Appl1

#### ALIGNMENTS

RESULT 1

US-09-966-880A-7

Sequence 7, Application US/09966880A

Patent No. US20020164743A1

GENERAL INFORMATION:

APPLICANT: Horijo, Taseuku

FILE OF INVENTION: NOVEL CYTIDINE DEAMINASE

FILE REFERENCE: 06501-088001

CURRENT APPLICATION NUMBER: US/09966, 880A

CURRENT FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: PCT/JP00/01918

PRIOR FILING DATE: 2000-03-28

PRIOR APPLICATION NUMBER: JP 11-371382

PRIOR FILING DATE: 1999-12-27

PRIOR APPLICATION NUMBER: JP 11-178999

PRIOR FILING DATE: 1999-06-24

PRIOR APPLICATION NUMBER: JP 11-87192

PRIOR FILING DATE: 1999-03-29

NUMBER OF SEQ ID NOS: 36

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 2818

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (80)...(673)

FEATURE:

NAME/KEY: 5'UTR

LOCATION: (1)...(79)

FEATURE:

NAME/KEY: 3'UTR

LOCATION: (677)...(2818)

US-09-966-880A-7

Query Match 100.0%; Score 2818; DB 9; Length 2818;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGAACCATCATTAATTAAGTGAAGATTTTCTGGCCGTGAGACTTGCAGAGAGGCAAGA 60  
Db 1 AGAGAACCATCATTAATTAAGTGAAGATTTTCTGGCCGTGAGACTTGCAGAGAGGCAAGA 60  
QY 61 AGACACTGTGACACCCCTATGACAGCCCTTGAAGAACCCGAGGAAGTTTCTTTACCA 120  
Db 61 AGACACTGTGACACCCCTATGACAGCCCTTGAAGAACCCGAGGAAGTTTCTTTACCA 120  
QY 121 ATTCAAAAATGTCGCGCTGAGGCTAAGGGTGCAGGCTGAGACCTACCTGTGCTAGTATGA 180  
Db 121 ATTCAAAAATGTCGCGCTGAGGCTAAGGGTGCAGGCTGAGACCTACCTGTGCTAGTATGA 180  
QY 181 GAGGCGTGAAGTGTCTATCTCTTTCTACTGACCTTGGTATCTTGGCAATPAGAACGG 240  
Db 181 GAGGCGTGAAGTGTCTATCTCTTTCTACTGACCTTGGTATCTTGGCAATPAGAACGG 240  
QY 241 CTGCGACGTGGAATGTCTCTTCCCTCCGCTACATCTGAGACTGGGACCTTGACCTTGGCCG 300  
Db 241 CTGCGACGTGGAATGTCTCTTCCCTCCGCTACATCTGAGACTGGGACCTTGACCTTGGCCG 300  
QY 301 CTGCTACCGGCTACCGCTGCTTCACTCTGAGAGCCCTGCTAGACTGAGACCTTGGCCG 360  
Db 301 CTGCTACCGGCTACCGCTGCTTCACTCTGAGAGCCCTGCTAGACTGAGACCTTGGCCG 360  
QY 361 GGCAGACTTCTGAGAGAGAACCCCAACTCACTGAGAGATCTTACCGCGGCTCTA 420  
Db 361 GGCAGACTTCTGAGAGAGAACCCCAACTCACTGAGAGATCTTACCGCGGCTCTA 420  
QY 421 CTCTGTGAGAGACCGCAAGGCTGAGCCCGAGGCGCTGCGCGGCTGACCGCGCGGCT 480  
Db 421 CTCTGTGAGAGACCGCAAGGCTGAGCCCGAGGCGCTGCGCGGCTGACCGCGCGGCT 480  
QY 481 GCAAAATGCAATGACCTTCAAAAGATATTTTTCCTGCTGAAATCTTTGTAGAAAA 540  
Db 481 GCAAAATGCAATGACCTTCAAAAGATATTTTTCCTGCTGAAATCTTTGTAGAAAA 540  
QY 541 CCATGAAAGAACTTTCAAAGCTGAGGAGGGCTGATGAAATTCAGTTCGCTCCAG 600  
Db 541 CCATGAAAGAACTTTCAAAGCTGAGGAGGGCTGATGAAATTCAGTTCGCTCCAG 600  
QY 601 ACAAGCTTGGCGGATCTTTGCCCCGTGATGAGGTTGATGACTTACGAGACGCAATTG 660  
Db 601 ACAAGCTTGGCGGATCTTTGCCCCGTGATGAGGTTGATGACTTACGAGACGCAATTG 660  
QY 661 TACTTTGGGACTTTGATAGCACTTCCAGGAATGTCAACAGATGAAATATCTGCTG 720  
Db 661 TACTTTGGGACTTTGATAGCACTTCCAGGAATGTCAACAGATGAAATATCTGCTG 720  
QY 721 AAGCAGTGTGATAAAGAGAGCTTCAAGTCTCTCTGTTTATCTCAACTCTGAC 780  
Db 721 AAGCAGTGTGATAAAGAGAGCTTCAAGTCTCTCTGTTTATCTCAACTCTGAC 780  
QY 781 TTTCTTGAAGTTTACGAAAAAATATTTATATAGACTCTTTTAAAAAGATCTATGCTG 840  
Db 781 TTTCTTGAAGTTTACGAAAAAATATTTATATAGACTCTTTTAAAAAGATCTATGCTG 840  
QY 841 AAAATAGAGAAAGAACAGAGCTGAGGAGAGGCTGCAATTTGTCAGTTTGAAT 900  
Db 841 AAAATAGAGAAAGAACAGAGCTGAGGAGAGGCTGCAATTTGTCAGTTTGAAT 900  
QY 901 GCAACATGTCTCCCTACTGGGAATTAACAGACTGCGAGACTTGGGAGCACTCTAAAGT 960  
Db 901 GCAACATGTCTCCCTACTGGGAATTAACAGACTGCGAGACTTGGGAGCACTCTAAAGT 960  
QY 961 CAACGTTTCTATGACTTTTATGATGAGTGAAGAGCAAGGATGATCTTAAAAAGCATG 1020  
Db 961 CAACGTTTCTATGACTTTTATGATGAGTGAAGAGCAAGGATGATCTTAAAAAGCATG 1020  
QY 1021 GTGAGAGATCAAAATGTTTATATCAACATCTTATATTTGATTCATTTGAGTTAAC 1080  
Db 1021 GTGAGAGATCAAAATGTTTATATCAACATCTTATATTTGATTCATTTGAGTTAAC 1080

QY 1081 AGTGCTGTATGATTAATTTTCTATATCTTTTCCCTGACGTTTACTTCAAGTAAAC 1140  
Db 1081 AGTGCTGTATGATTAATTTTCTATATCTTTTCCCTGACGTTTACTTCAAGTAAAC 1140  
QY 1141 AAACCTTCCATGAGGCAATGATCTATAGAACCTCCCTAATGAGAGATCTGGGATGTG 1200  
Db 1141 AAACCTTCCATGAGGCAATGATCTATAGAACCTCCCTAATGAGAGATCTGGGATGTG 1200  
QY 1201 GACCCAAACCATCTCTCAAGCATTAATATCAATCATGCGCTGATGTTTAAATCAG 1260  
Db 1201 GACCCAAACCATCTCTCAAGCATTAATATCAATCATGCGCTGATGTTTAAATCAG 1260  
QY 1261 CAGAGCATGTTTATGTTTGTACAAAGAAAGATTGTTATGGGTGGGATGAGGTAT 1320  
Db 1261 CAGAGCATGTTTATGTTTGTACAAAGAAAGATTGTTATGGGTGGGATGAGGTAT 1320  
QY 1321 GACCATGATGCTACCTTGAAGCTACTTAATPAGAACTTAAATGGGAGAGGAG 1380  
Db 1321 GACCATGATGCTACCTTGAAGCTACTTAATPAGAACTTAAATGGGAGAGGAG 1380  
QY 1381 TGTGAACAAGACACCTTAATATGAGTGTCTGATGAGAGAAATCTTCTGAAACG 1440  
Db 1381 TGTGAACAAGACACCTTAATATGAGTGTCTGATGAGAGAAATCTTCTGAAACG 1440  
QY 1441 AAACCTTTTAAAGAACTCCCTAATATTAAGAAACCCCACTTCAATATCATATTA 1500  
Db 1441 AAACCTTTTAAAGAACTCCCTAATATTAAGAAACCCCACTTCAATATCATATTA 1500  
QY 1501 GCAAACTTGAAGAAAGTGTCTTGAATGTTGGGAGAGAAAAATCTATGCTCTCGT 1560  
Db 1501 GCAAACTTGAAGAAAGTGTCTTGAATGTTGGGAGAGAAAAATCTATGCTCTCGT 1560  
QY 1561 GGGTCTGTATCTCAGAAATGCAATCAGGTCAAGTGTCTGACATTTGTATGTTGT 1620  
Db 1561 GGGTCTGTATCTCAGAAATGCAATCAGGTCAAGTGTCTGACATTTGTATGTTGT 1620  
QY 1621 GATGCTTCTCCAAAGTATATTAATTAAGAGAGTTGTGACAAAACAGATATTA 1680  
Db 1621 GATGCTTCTCCAAAGTATATTAATTAAGAGAGTTGTGACAAAACAGATATTA 1680  
QY 1681 AGCTGCAACCGTGGCAACGCTCATGTTCTGAGCTTGGAGGTTGAGAGGAGGA 1740  
Db 1681 AGCTGCAACCGTGGCAACGCTCATGTTCTGAGCTTGGAGGTTGAGAGGAGGA 1740  
QY 1741 TGGCTTGAACACAGGTGTTCAAGGCGAGCTTGGGCAACATTAAGATCTGTCTCA 1800  
Db 1741 TGGCTTGAACACAGGTGTTCAAGGCGAGCTTGGGCAACATTAAGATCTGTCTCA 1800  
QY 1801 AAAAAAAGAAAAAAGAAAGAGAGGAGGCGGCTGAGTCAAGCTGTATCC 1860  
Db 1801 AAAAAAAGAAAAAAGAAAGAGAGGAGGCGGCTGAGTCAAGCTGTATCC 1860  
QY 1861 CAGCACTTGGAGGCGGAGCGGCGGATCACTGTGTGATCAGAGGTTTGAACAAGCT 1920  
Db 1861 CAGCACTTGGAGGCGGAGCGGCGGATCACTGTGTGATCAGAGGTTTGAACAAGCT 1920  
QY 1921 GGCCTAATGAGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAGGCTGTAGC 1980  
Db 1921 GGCCTAATGAGCAAAACCCGCTGTATCTCAAAATGCAAAATTAAGCAGGCTGTAGC 1980  
QY 1981 AGGCACTGTATCCGACTACTTGGAGGCTGAGGAGAGGATGCTTGAACCAAGA 2040  
Db 1981 AGGCACTGTATCCGACTACTTGGAGGCTGAGGAGAGGATGCTTGAACCAAGA 2040  
QY 2041 GGTGAGGTTGCAAGTAACTGATGCTGCTGCTGCACTCAGCTTGGGAGCAAGAGCA 2100  
Db 2041 GGTGAGGTTGCAAGTAACTGATGCTGCTGCTGCACTCAGCTTGGGAGCAAGAGCA 2100  
QY 2101 AGACTTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAAATATTTG 2160  
Db 2101 AGACTTGTCTCAGAAAAAAGAAAAAAGAGAGAGAGAGAGAAATATTTG 2160

QY 2161 GGAGAGAGATGGGAGAGATTCAGAGAAATTTGCTTTATTCAGAGAAATGTAGGA 2220  
DB 2161 GGAGAGAGATGGGAGAGATTCAGAGAAATTTGCTTTATTCAGAGAAATGTAGGA 2220  
QY 2221 GCCAATAGAGGATCCCTATTGCTCTTTGCTGCTATTGCTCCCTAACTGCTCTT 2280  
DB 2221 GCCAATAGAGGATCCCTATTGCTCTTTGCTGCTATTGCTCCCTAACTGCTCTT 2280  
QY 2281 GACAGTGAAGAAATATTCAGATTAACCATATCCCTGCTGCTATTGCTCCCTAACTGCTCTT 2340  
DB 2281 GACAGTGAAGAAATATTCAGATTAACCATATCCCTGCTGCTATTGCTCCCTAACTGCTCTT 2340  
QY 2341 TGCAATGAGATGAGAGATTCAGAGAAATTTGCTTTATTCAGAGAAATGTAGGA 2400  
DB 2341 TGCAATGAGATGAGAGATTCAGAGAAATTTGCTTTATTCAGAGAAATGTAGGA 2400  
QY 2401 TTATGTACATTAAGTTGTAAAGAGTTAAAGATTTGCTTACCTGATGATTTATATAT 2460  
DB 2401 TTATGTACATTAAGTTGTAAAGAGTTAAAGATTTGCTTACCTGATGATTTATATAT 2460  
QY 2461 TTATATATTTTTCGCTTAATGATTTTATTAACATGATTTCTTTGCTGATATATGA 2520  
DB 2461 TTATATATTTTTCGCTTAATGATTTTATTAACATGATTTCTTTGCTGATATATGA 2520  
QY 2521 AATGAGTCTCAAGCTTCATTAATTTATTAACCTTTGAATGATTTATTAACAAGTAT 2580  
DB 2521 AATGAGTCTCAAGCTTCATTAATTTATTAACCTTTGAATGATTTATTAACAAGTAT 2580  
QY 2581 GTATTTGATACATTTGAGATGATGCTGCTGAGAGCAATTTCTTTGATTTTATTAAGTAT 2640  
DB 2581 GTATTTGATACATTTGAGATGATGCTGCTGAGAGCAATTTCTTTGATTTTATTAAGTAT 2640  
QY 2641 TTATGACAGCAATTTGCTTCTGCTCACTTCAATGATTTAAATTAAGTATTAATAT 2700  
DB 2641 TTATGACAGCAATTTGCTTCTGCTCACTTCAATGATTTAAATTAAGTATTAATAT 2700  
QY 2701 TTGGAAGCTGTAAGATTAATTAACCAATTAATTAATTAATTAATTAATTAAT 2760  
DB 2701 TTGGAAGCTGTAAGATTAATTAACCAATTAATTAATTAATTAATTAATTAAT 2760  
QY 2761 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2818  
DB 2761 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2818

RESULT 2  
US-09-966-880a-10  
Sequence 10, Application US/09966880A  
Patent No. US20020164743A1  
GENERAL INFORMATION:  
APPLICANT: Honjo, Taku  
APPLICANT: Muramatsu, Masamichi  
TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE  
FILE REFERENCE: 06501-088001  
CURRENT APPLICATION NUMBER: US/09/966,880A  
PRIOR APPLICATION NUMBER: PCT/JP00/01918  
PRIOR FILING DATE: 2000-03-28  
PRIOR APPLICATION NUMBER: JP 11-371382  
PRIOR FILING DATE: 1999-12-27  
PRIOR APPLICATION NUMBER: JP 11-178999  
PRIOR FILING DATE: 1999-06-24  
PRIOR APPLICATION NUMBER: JP 11-87192  
PRIOR FILING DATE: 1999-03-29  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 6564  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-966-880a-10

Query Match 77.2%; Score 2174.6; DB 9; Length 6564;

Beat Local Similarity 99.8%; Pred. No. 0;  
Matches 2177; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 617 CTTTGGCCCTGATGAGGTTGATGATCTTACGAGACCATTTGCTTGGGACTTTGA 676  
DB 3735 CTACAGCCCTGATGAGGTTGATGATCTTACGAGACCATTTGCTTGGGACTTTGA 3794  
QY 677 TAGCACTTCCAGGAATGTCACACAGATGAATATCTTCTGGAAGACGTGATAAA 736  
DB 3795 TAGCACTTCCAGGAATGTCACACAGATGAATATCTTCTGGAAGACGTGATAAA 3854  
QY 737 AACGCTTCAAGCTCTCTGTTTATTTCTTCACTCTCACTTTCTTGAAGTTTACA 796  
DB 3855 AACGCTTCAAGCTCTCTGTTTATTTCTTCACTCTCACTTTCTTGAAGTTTACA 3914  
QY 797 GAAAAATATTTATATATGCACTCTTAAAGATCTATGCTTGAATATGAGAGAAC 856  
DB 3915 GAAAAATATTTATATATGCACTCTTAAAGATCTATGCTTGAATATGAGAGAAC 3974  
QY 857 ACAGGCTGGCCAGGAGCGTGTGCAATGATGAGATTTGAATGCAATGCTCCCTA 916  
DB 3975 ACAGGCTGGCCAGGAGCGTGTGCAATGATGAGATTTGAATGCAATGCTCCCTA 4034  
QY 917 CTGGGAATTAAGAACTGCAAGACCTTGGGACATCTTAAAGTCAACGTTTCTATGA 976  
DB 4035 CTGGGAATTAAGAACTGCAAGACCTTGGGACATCTTAAAGTCAACGTTTCTATGA 4094  
QY 977 CTTTGAAGTGAAGAGAGAGAGAGATGATCTTAAAGATGAGAGAGATGCAATG 1036  
DB 4095 CTTTGAAGTGAAGAGAGAGAGATGATCTTAAAGATGAGAGAGATGCAATG 4154  
QY 1037 TTTTATATCAACATCTTATTTATTTGATTTCAATTTGAGTAAACAGTGAGTATGATA 1096  
DB 4155 TTTTATATCAACATCTTATTTATTTGATTTCAATTTGAGTAAACAGTGAGTATGATA 4214  
QY 1097 GATTTTCTATCTTTTCCCTTGAAGTTTACTTCAAGTAAACAACTCTTCATCAGG 1156  
DB 4215 GATTTTCTATCTTTTCCCTTGAAGTTTACTTCAAGTAAACAACTCTTCATCAGG 4274  
QY 1157 CCATGATCTATAGACCTCTTATGAGAGATCTGGGTGATTTGACCCCAACATCTC 1216  
DB 4275 CCATGATCTATAGACCTCTTATGAGAGATCTGGGTGATTTGACCCCAACATCTC 4334  
QY 1217 TCCAAAGCTTAATATCAATCATGCGCTGATGTTTATTCAGAGAGATGTTTGA 1276  
DB 4335 TCCAAAGCTTAATATCAATCATGCGCTGATGTTTATTCAGAGAGATGTTTGA 4394  
QY 1277 TGTTTGACAAAAGAGATTTGATGGTGGGATGAGATATAGACATGATGCTCAC 1336  
DB 4395 TGTTTGACAAAAGAGATTTGATGGTGGGATGAGATATAGACATGATGCTCAC 4454  
QY 1337 CTTCAAGCTTCTTAAATGAATCTTAAATGGGAGAGAGACCTGAAACAAGACCC 1396  
DB 4455 CTTCAAGCTTCTTAAATGAATCTTAAATGGGAGAGAGACCTGAAACAAGACCC 4514  
QY 1397 TAAATAGGTTGATGCTGAGATGATGATCTTGGAAACCAACCTTTTAAAGAA 1456  
DB 4515 TAAATAGGTTGATGCTGAGATGATGATCTTGGAAACCAACCTTTTAAAGAA 4574  
QY 1457 GTCCCTATTTTAAACAACCAACCACTTCAATCATATATAGCAAACTTGAAGG 1516  
DB 4575 GTCCCTATTTTAAACAACCAACCACTTCAATCATATATAGCAAACTTGAAGG 4634  
QY 1517 AAGTGTGATATGTTGGGAGAGAAATCTATGCTGCTGCTGCTCTTCAATCTCA 1576  
DB 4635 AAGTGTGATATGTTGGGAGAGAAATCTATGCTGCTGCTGCTCTTCAATCTCA 4694  
QY 1577 GAAATGCAATCAAGTCAAGTCTTCACTTTTGAATGCTGCTTCTTCCCAAG 1636  
DB 4695 GAAATGCAATCAAGTCAAGTCTTCACTTTTGAATGCTGCTTCTTCCCAAG 4754  
QY 1637 GTATATTAATATATAGAGATTTGACAAAACAAATGATTAAGCTGGAACGCTGGC 1696



QY	1157	CCATGATCTATAGGACCTCCATAATGAGAAGATCTGGGTGATTGGACCCCAACCATCTTC	1216
Db	9490	CCATGATCTATAGGACCTCCATAATGAGAAGATCTGGGTGATTGGACCCCAACCATCTTC	9549
QY	1217	TCCAAAGCATTAATATCCAAATCATGCGCTGATGTTTAAATCAGCAGAAAGATGTTTAA	1276
Db	9550	TCCAAAGCATTAATATCCAAATCATGCGCTGATGTTTAAATCAGCAGAAAGATGTTTAA	9609
QY	1277	TGTTTGTATCAAAAAGATTTGTTATTTGGGTGGGATGAGGTATATGACATGATGGTCAC	1336
Db	9610	TGTTTGTATCAAAAAGATTTGTTATTTGGGTGGGATGAGGTATATGACATGATGGTCAC	9669
QY	1337	CTTCAGCTACTTTAATATAAAGATCTTAAATGCGCAGAGGACCTGTGAAACAACCC	1396
Db	9670	CTTCAGCTACTTTAATATAAAGATCTTAAATGCGCAGAGGACCTGTGAAACAACCC	9729
QY	1397	TAAATATGGGTGATGTCGTGAGTAGCAAACTCTTGTGAAACGCAACTCTTTAAGAA	1456
Db	9730	TAAATATGGGTGATGTCGTGAGTAGCAAACTCTTGTGAAACGCAACTCTTTAAGAA	9789
QY	1457	GTCCCTAATTTAGAAACCCACAACCTTCAATATCATATATTAAGCAACAATGGAAG	1516
Db	9790	GTCCCTAATTTAGAAACCCACAACCTTCAATATCATATATTAAGCAACAATGGAAG	9849
QY	1517	AAGTTGCTGTGAATGTTGGGAGAGGAAATCTATGTCCTCGTGGGTCTCTTCACTCA	1576
Db	9850	AAGTTGCTGTGAATGTTGGGAGAGGAAATCTATGTCCTCGTGGGTCTCTTCACTCA	9909
QY	1577	GAAATGCAATCAGGTCAGAGTTTGCTATCAATTTTGTATGTGTGATGCTTCCCAAG	1636
Db	9910	GAAATGCAATCAGGTCAGAGTTTGCTATCAATTTTGTATGTGTGATGCTTCCCAAG	9969
QY	1637	GTAATATTAACTATATTAAGAGATTGTGACAAAACAGATGATTAAGCTGCGAACCTGGC	1696
Db	9970	GTAATATTAACTATATTAAGAGATTGTGACAAAACAGATGATTAAGCTGCGAACCTGGC	10022
QY	1697	ACAGCTCATATGTTCTATGCTGCTTGGGAGGTTGAGAGGAGAGATGGCTTGAACAAGT	1756
Db	10030	ACAGCTCATATGTTCTATGCTGCTTGGGAGGTTGAGAGGAGAGATGGCTTGAACAAGT	10083
QY	1757	GTTCAAGGCGACGCTGGGCAACATTAACAATCTGTCTCTCAAAAAAATTTTTTTTTT	1816
Db	10090	GTTCAAGGCGACGCTGGGCAACATTAACAATCTGTCTCTCAAAAAAATTTTTTTTTT	10144
QY	1817	AAGAAAGAGAGGCGCGCGCGTGTGGCTCAGCGCTGTATCCAGCACTTTGGGAGCG	1876
Db	10150	AAGAAAGAGAGGCGCGCGCGTGTGGCTCAGCGCTGTATCCAGCACTTTGGGAGCG	10203
QY	1877	CGAGCCGGGCGGATCACCTGTGGTCAGAGGTTTGACACACCTGGCCAAATGCGCAAA	1936
Db	10210	CGAGCCGGGCGGATCACCTGTGGTCAGAGGTTTGACACACCTGGCCAAATGCGCAAA	10263
QY	1937	CCCGCTGTATCTCAAAAATGCAAAAATTAGCCAGCGGTGGTACAGGCACCTGTATCCC	1996
Db	10270	CCCGCTGTATCTCAAAAATGCAAAAATTAGCCAGCGGTGGTACAGGCACCTGTATCCC	10323
QY	1997	AGTACTTTGGGAGGCTGAGCAGAGGATGCTTGAACCCAGAGGTTGAGGTTTGACGTA	2056
Db	10330	AGTACTTTGGGAGGCTGAGCAGAGGATGCTTGAACCCAGAGGTTGAGGTTTGACGTA	10383
QY	2057	AGCTGAATGTCGCTGTTGCACTTCACGCTGGCGCACAAGACACAACTCTGTCTGAA	2116
Db	10390	AGCTGAATGTCGCTGTTGCACTTCACGCTGGCGCACAAGACACAACTCTGTCTGAA	10444
QY	2117	AAAAAATTTTTAAGAGAGAGAGGAGAAAAGAACTATTTTGGAGAGAAAGATGGGG	2176
Db	10450	AAAAAATTTTTAAGAGAGAGAGGAGAAAAGAACTATTTTGGAGAGAAAGATGGGG	10503
QY	2177	AAGCATTTGCAAGAAATTTGCTTTATCTCAACAAATGTATAGAGCCATATAGGATCCC	2236
Db	10510	AAGCATTTGCAAGAAATTTGCTTTATCTCAACAAATGTATAGAGCCATATAGGATCCC	10563
QY	2237	TATTTGTCTCTTTTGTGTCTAATTTGTCCCTAACCATGTCTTTGACATGAGAAAAATA	2296

Db	10570	TATTGCTCTTTTGGATGCTATTTTCCCTTAACAACCTGCTTTTGACGTGAGAAAAATA	10623
Qy	2297	TTGAGAAATACATATCCCTGCGGCTTATTAACAGCAACCCCTTGCAATGAAGTGAAC	2356
Db	10630	TTGAGAAATACATATCCCTGCGGCTTATTAACAGCAACCCCTTGCAATGAAGTGAAC	10689
Qy	2257	AGATCCACAGGAAACCTTGAAATGCAACAATGCTTTTAAATCTTATTTGACATTAAGT	2416
Db	10690	AGATCCACAGGAAACCTTGAAATGCAACAATGCTTTTAAATCTTATTTGACATTAAGT	10749
Qy	2417	TGTAAAGAGTTAAATGTTACTCATGATTTCAATTAATATTTATTTTGGCT	2476
Db	10750	TGTAAAGAGTTAAATGTTACTCATGATTTCAATTAATATTTATTTTGGCT	10805
Qy	2477	CTAATGATTTTATTTATTAACATGATTTCCCTTTCTGATATATTGAATGAGTCTCAAGC	2536
Db	10810	CTAATGATTTTATTTATTAACATGATTTCCCTTTCTGATATATTGAATGAGTCTCAAGC	10866
Qy	2537	TTCAATTAATTTAATACCTTTAGAAATGATTTCTAATAACACGATGATATTGTAACATTGC	2596
Db	10870	TTCAATTAATTTAATACCTTTAGAAATGATTTCTAATAACACGATGATATTGTAACATTGC	10922
Qy	2597	AGTAAAGTGTCTGAGGACCATTCCTCTGATTTTATTAATTAATGACAGCAATT	2656
Db	10930	AGTAAAGTGTCTGAGGACCATTCCTCTGATTTTATTAATTAATGACAGCAATT	10989
Qy	2657	TGCTTGAGCTCATCTTTCATCATGATTAATTAATGATTAATTTTGGACCTGTGAAG	2716
Db	10990	TGCTTGAGCTCATCTTTCATCATGATTAATTAATGATTAATTTTGGACCTGTGAAG	11049
Qy	2717	ATAAATACCAATTAATTAATTAATTAATGATTTATATGAAGTTAAATTAATTAATCAG	2776
Db	11050	ATAAATACCAATTAATTAATTAATTAATGATTTATATGAAGTTAAATTAATTAATCAG	11109
Qy	2777	TATGATGGAATTAACCTGAAA 2797	
Db	11110	TATGATGGAATTAACCTGAGA 11130	

RESULT 4

US-09-966-880A-15

Sequence 15, Application US/09966880A

Patent No. US20020164743A1

GENERAL INFORMATION:

APPLICANT: Honjo, Tasuku

APPLICANT: Muramatsu, Masamichi

TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE

FILE REFERENCE: 06501-088001

CURRENT APPLICATION NUMBER: US/09/966, 880A

CURRENT FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: PCT/JP00/01918

PRIOR FILING DATE: 2000-03-28

PRIOR APPLICATION NUMBER: JP 11-371382

PRIOR FILING DATE: 1999-12-27

PRIOR APPLICATION NUMBER: JP 11-178999

PRIOR FILING DATE: 1999-06-24

PRIOR APPLICATION NUMBER: JP 11-87192

PRIOR FILING DATE: 1999-03-29

NUMBER OF SEQ ID NOS: 36

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15

LENGTH: 2172

TYPE: DNA

ORGANISM: Homo sapiens

US-09-966-880A-15

Query Match 77.1%; Score 2172; DB 9; Length 2172;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

623 CCCCTGTATGAGTTATGATGACTTACAGACGCAATTCGATCTTTGGAGCTTTGATATGCA 682

Db	CCCCGTATGAGGTGATGACCTTACGAGACGCAATTTGCTACTTTGGGACTTTGACGA	60
Qy	CTTCCGAGATGTCAACAACGATGAAATATCTTCGTGAGACAGTGGATAAAAACAGT	742
Db	CTTCCGAGGATGTCAACAACGATGAAATATCTTCGTGAGACAGTGGATAAAAACAGT	120
Qy	CCCTCAAGCTCTCTGTTTTATCTTCAACTCTCACTTCTTACAGTTACAGAAAA	802
Db	121 CCTTCAAGTCTCTCTGTTTTATCTTCAACTCTCACTTCTTACAGTTACAGAAAA	180
Qy	ATATTATATACGACTTTAAAAAGCTATGCTTGAAATAGAAAGAAACAAGGT	862
Db	181 ATATTATATACGACTTTAAAAAGCTATGCTTGAAATAGAAAGAAACAAGGT	240
Qy	863 CTGGCCAGGGACGTGCTGCATTTGGTGCAGTTTGAATGCAACTGTCCCTTACGGGA	922
Db	241 CTGGCCAGGGACGTGCTGCATTTGGTGCAGTTTGAATGCAACTGTCCCTTACGGGA	300
Qy	923 ATTAACGAACCTGCAGACCTGGGACATCTTAAGTGTCAACGTTTTCTATACCTTTA	982
Db	301 ATTAACGAACCTGCAGACCTGGGACATCTTAAGTGTCAACGTTTTCTATACCTTTA	360
Qy	983 GGTAGATAGAGACGAGAGTGAATCTTAAAAAGCATGTAAGAGATCAATGTTTTA	1044
Db	361 GGTAGATAGAGACGAGAGTGAATCTTAAAAAGCATGTAAGAGATCAATGTTTTA	420
Qy	1043 TATCAACATCTTTATTTATTTGATTCAATTGAGTTAACAGGTGTGATGATGATTT	1108
Db	421 TATCAACATCTTTATTTATTTGATTCAATTGAGTTAACAGGTGTGATGATGATTT	480
Qy	1103 TCTATTCTTTCCCTTGAGCTTACTTTCAAGTACACAAACTTTGCATCAGGCCATGA	1168
Db	481 TCTATTCTTTCCCTTGAGCTTACTTTCAAGTACACAAACTTTGCATCAGGCCATGA	540
Qy	1163 TCTATTGAGACCTCTTAATGAGAGTACTGGGAGATTGGAACCCAAACATCTCTCAAA	1222
Db	541 TCTATTGAGACCTCTTAATGAGAGTACTGGGAGATTGGAACCCAAACATCTCTCAAA	600
Qy	1223 GCATTATATCCCATCATGCGCTGTATGTTTTAACTACAGCAAGACATGTTTTATGTTG	1282
Db	601 GCATTATATCCCATCATGCGCTGTATGTTTTAACTACAGCAAGACATGTTTTATGTTG	660
Qy	1283 TACAAAAGAAATGTTATGGGTGGGAGTGAAGTATAGACATGCAATGCTCACTTCAA	1342
Db	661 TACAAAAGAAATGTTATGGGTGGGAGTGAAGTATAGACATGCAATGCTCACTTCAA	720
Qy	1343 GCATCTTAAATAAAGATCTTAAATGGGACAGAGGACTGGAACAAGACACCTTAATAA	1402
Db	721 GCATCTTAAATAAAGATCTTAAATGGGACAGAGGACTGGAACAAGACACCTTAATAA	780
Qy	1403 TGGGTTGATGCTGGAATAGCAAAATCTTCTGGAAACGCAACTTTTAAAGAAATCCCT	1462
Db	781 TGGGTTGATGCTGGAATAGCAAAATCTTCTGGAAACGCAACTTTTAAAGAAATCCCT	840
Qy	1463 AATTTCGAAAACACCCCAAACTTCACATATCAATAATTAGCAAAACAATTGAAAGAAATG	1522
Db	841 AATTTCGAAAACACCCCAAACTTCACATATCAATAATTAGCAAAACAATTGAAAGAAATG	900
Qy	1523 CTGGAATGTTGGGGAAGAGAAATCTATTTGGCTCGGGGTCCTTCATCTCAGAAATG	1582
Db	901 CTGGAATGTTGGGGAAGAGAAATCTATTTGGCTCGGGGTCCTTCATCTCAGAAATG	960
Qy	1583 CCATCAGGTCAAGGTTTGCTCATTTTGTATGTTGTGATGCTTCTCCAAAGGTATAT	1644
Db	961 CCATCAGGTCAAGGTTTGCTCATTTTGTATGTTGTGATGCTTCTCCAAAGGTATAT	1022
Qy	1643 TAACTATATAGAGAGTTGTGACAAACAAGATGATAAAGCTGCGAACCTGTGCACACGC	1702
Db	1021 TAACTATATAGAGAGTTGTGACAAACAAGATGATAAAGCTGCGAACCTGTGCACACGC	1082
Qy	1703 TCAATAGTCTAAGCTGTTGGAGAGTTGAGAGAGGAGATGGCTTGAACAAGGTGTTCAA	1762
Db	1081 TCAATAGTCTAAGCTGTTGGAGAGTTGAGAGAGGAGATGGCTTGAACAAGGTGTTCAA	1144

OY	1763	GGCCAGCCTGGGGCAACATTAACAAGATCTCTGCTCTCAAAAAAAAAAAAAAAAAAAAAAGAA	1822
Db	1141	GGCCAGCCTGGGGCAACATTAACAGATCTCTGCTCTCAAAAAAAAAAAAAAAAAAAAAAGAA	1200
OY	1823	GAGAGAGAGCGCGGGCGTGTGGCTCAGCGCCTGTATATCCAGCACTTTGGAGGCGAGCC	1882
Db	1201	GAGAGAGAGCGCGGGCGTGTGGCTCAGCGCCTGTATATCCAGCACTTTGGAGGCGAGCC	1260
OY	1883	GGGCGGATCACCTGTGGTCAGAGTTTGAACCCAGCCTGGCCAAATGGCMAAACCCCGT	1942
Db	1261	GGGCGGATCACCTGTGGTCAGAGTTTGAACCCAGCCTGGCCAAATGGGCAAAACCCCGT	1320
OY	1343	CTGTAACTCAAAATGCAAAAAATTAGCCAGGGGTGGTACAGAGCACCTGTATATCCAGCTAC	2002
Db	1321	CTGTAACTCAAAATGCAAAAAATTAGCCAGGGGTGGTACAGAGCACCTGTATATCCAGCTAC	1380
OY	2003	TTGGGAGGCTGAGGCAAGAAATGCTTTGAACCCAGAGGCTGAGGTTTGGAGTAAGCTGA	2062
Db	1381	TTGGGAGGCTGAGGCAAGAAATGCTTTGAACCCAGAGGCTGAGGTTTGGAGTAAGCTGA	1440
OY	2063	GATGTCGCCGTTCGACTCCAGCCTGGCGCAACAAGACAAGACTCTGTCTCAGAAAAAAA	2122
Db	1441	GATGTCGCCGTTCGACTCCAGCCTGGCGCAACAAGACAAGACTCTGTCTCAGAAAAAAA	1500
OY	2123	AAAAAAAAAGAGAGAGAGAGAAAGAAACAATTTGGAGAGAGAGATGGGGACACT	2182
Db	1501	AAAAAAAAAGAGAGAGAGAGAAAGAAACAATTTGGAGAGAGAGATGGGGACACT	1560
OY	2183	TGCAAGAGAAATTTGCTTTATCCAACAAAATGTATAGAGCAATTAAGGATCCCTATTGG	2242
Db	1561	TGCAAGAGAAATTTGCTTTATCCAACAAAATGTATAGAGCAATTAAGGATCCCTATTGG	1620
OY	2243	TCTCTTTTGGTGTCTATTGTTCCTTAACAACGTGCTTTGACAGTGAGAAAAATATTCAGA	2302
Db	1621	TCTCTTTTGGTGTCTATTGTTCCTTAACAACGTGCTTTGACAGTGAGAAAAATATTCAGA	1680
OY	2303	ATAACCAATTCCTGTGCCCTTATTAACCTAGCAACCTTGCAATGAATGAGCAATCC	2362
Db	1681	ATAACCAATTCCTGTGCCCTTATTAACCTAGCAACCTTGCAATGAATGAGCAAGATCC	1740
OY	2363	ACAGAGAAACCTGAATGCACAACTGTCTTATTTAACTTAATGTATATTAAGTTGTAA	2422
Db	1741	ACAGAGAAACCTGAATGCACAACTGTCTTATTTAACTTAATGTATATTAAGTTGTAA	1800
OY	2423	AGAGTTAAAAATGTACTTCATGTATTCATTTATATTTATATTTATTTGGCGCTAATG	2482
Db	1801	AGAGTTAAAAATGTACTTCATGTATTCATTTATATTTATATTTATTTGGCGCTAATG	1860
OY	2483	ATTTTATATTAACATGATTTCTTTTCTGTATATATGAATGAGAGTCAAAAGTTCAATA	2542
Db	1861	ATTTTATATTAACATGATTTCTTTTCTGTATATATGAATGAGAGTCAAAAGTTCAATA	1920
OY	2543	AATTTATTAACCTTNGAATATGTTTATAACAAGCATGTAAATTGTAAACATGTAGAAAT	2602
Db	1921	AATTTATTAACCTTNGAATATGTTTATAACAAGCATGTAAATTGTAAACATGTAGAAAT	1980
OY	2603	GGTGTCAAGAAAGCAATTTCTCTGATTTTATAGTAACTTTTATGACAGCAATTTGGCTC	2662
Db	1981	GGTGTCAAGAAAGCAATTTCTCTGATTTTATAGTAACTTTTATGACAGCAATTTGGCTC	2040
OY	2663	TGCGTCACTTTCATCAGTTAAATTAATGAATTAATTTTGGAGCGTGTAGAGATAAA	2722
Db	2041	TGCGTCACTTTCATCAGTTAAATTAATGAATTAATTTTGGAGCGTGTAGAGATAAA	2100
OY	2723	TACCAAAATAAATATATTTAAAGATTTATATGAAGTTAAATTAATAAATCACTATGAT	2782
Db	2101	TACCAAAATAAATATATTTAAAGATTTATATGAAGTTAAATTAATAAATCACTATGAT	2160
OY	2783	GGATTAACCTTG 2794	
Db	2161	GGATTAACCTTG 2172	

```
RESULT 5
US-09-966-880a-1
; Sequence 1, Application US/09966880A
; Patent No. US20020164743A1
; GENERAL INFORMATION:
; APPLICANT: Honjo, Tasuku
; APPLICANT: Muramatsu, Masamichi
; TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
; FILE REFERENCE: 06501-088001
; CURRENT APPLICATION NUMBER: US/09/966,880A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: PCT/JP00/01918
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 11-371382
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: JP 11-178999
; PRIOR FILING DATE: 1999-06-24
; PRIOR APPLICATION NUMBER: JP 11-871992
; PRIOR FILING DATE: 1999-03-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2440
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (93)...(686)
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(92)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (690)...(2440)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2440)
; OTHER INFORMATION: n = A,T,C or G
US-09-966-880a-1
```

```
Query Match 21.4%; Score 603.4; DB 9; Length 2440;
Best Local Similarity 69.4%; Pred. No. 4e-118;
Matches 1006; Conservative 0; Mismatches 351; Indels 92; Gaps 10;
```

```
QY 4 GAACCATCATTATTAAGTGAAGATTTTCTGGCCGAGACTTGACGGAGGCAAG-AAG 62
DB 16 GAAGCAAGCTTGCTGAAGCAAGCTTCTGGCTTAAGACTTTGAGGAGTCAAGAAAG 75
QY 63 ACACTCTGACACACACTATGACAGCCTCTTGATGACCGAGAAAGTTCTTTACAA 122
DB 76 TCACGGCTGAGAGACCATATGAGACAGCTTCTGATGAGCAAAAGAGTTCTTTACAT 135
QY 123 TCAAAATGTCGCTGAGCTAAGAGGTGCGGCTGAGAGCTTCTGCTAGTGAAGA 182
DB 136 TCAAAATGTCGCTGAGCTAAGAGGTGCGGCTGAGAGCTTCTGCTAGTGAAGA 195
QY 183 GGGGTGACGTGTACATCTTTTCACTGAGCTTTGATGCTTGGCAATAGAAAGCT 242
DB 196 GGAAGATATGTGCACTCTCTGCTCACTGAGACTTCGGCCACCTTGGCAACAAAGTGGCT 255
QY 243 GCCACGTGAAATGCTCTTCTCCGCTACATCTGAGACTGGAGCTGAGCCGCT 302
DB 256 GCCACGTGAAATGCTCTTCTCCGCTACATCTGAGACTGGAGCTGAGCCGCT 315
QY 303 GCTACCGGCTACCTGCTTCACTCTGAGCCCTGCTACGACTGTGCCGACATGTGG 362
DB 316 GTTACCGGCTACCTGCTTCACTCTGAGCCCTGCTACGACTGTGCCGACATGTGG 375
QY 363 CCGACTTTCTGAGAGGAAACCCCAAGCTCAAGTGTAGAGATCTTACCGGCGCTCTACT 422
DB 376 CTGAGTTCTGAGATGAACCCCAAGCTCAAGTGTAGAGATCTTACCGGCGCTCTACT 435
```

```
QY 423 TCTGTGAGAACCGGAAGCTGAGCCCGAGGGGCTGCGCGCTGCACCCGCGGGCTGC 482
DB 436 TCTGTGAAGACCGGAAGCTGAGCCCGAGGGGCTGCGCGCTGCACCCGCGGGCTGC 495
QY 483 AATAGCCATCATGACCTCAAGATTAATTTTACTGCTGAATATCTTTGTAGAAAAC 542
DB 496 AGATCGGATATATACCTTCAAGACATATTTTACTGCTGAATATCTTTGTAGAAAAC 555
QY 543 ATGAAAGACTTTCAAAAGCTGGGAAGGCTGATGAAAATTCAGTTGCTCTCCAGAC 602
DB 556 GTGAAAGACTTTCAAAAGCTGGGAAGGCTGATGAAAATTCAGTTGCTCTCCAGAC 615
QY 603 AGCTTGGCGGATCTTTTGGCCCTGATGAGGTTGATGACTTACGAGACGCAATTCGTA 662
DB 616 AACTTGGCGGATCTTTTGGCCCTGATGAGGTTGATGACTTACGAGACGCAATTCGTA 675
QY 663 CTTTGGGACTTGTATGACACTTCCAGGAATGTACACAGATGAAATATCTGCTGAA 722
DB 676 TGTGGGATTTTGAAGAACCTCTGGAATGTACACAGATGAAATATCTGCTGAA 732
QY 723 GACAGTGATTAATAAAGCTTCAAGTCTTCTGTTTATTTCTTCAACTCTCACTT 782
DB 733 GAGACTGATTAATAAAGCTTCAAGTCTTCTGTTTATTTCTTCAACTCTCACTT 790
QY 783 TCTTAGAGTTTACGAAAAAATATTTATATACACTCTTTAAAAAGATCATGTCTTGA 842
DB 791 TTATAGTGTAGGGGAAA-----TTATATGACTTTTAAAAAATATCTGAGCTGCAC 843
QY 843 AATGAGAGAGAAACAGAGTCTGCGCCAGGACGCTGCAATGTGCAATTTGAAATGC 902
DB 844 AGGACCGCAGAGGAATATGTAC-----TGAAGTCTGCTGTGC 882
QY 903 AACATGTCCTCTACT-GGGAAATPACAGAACTGAGAGACTGGAGAGATCTAAAGTGC 961
DB 883 AACATGTCCTCTACT-GGGAAATPACAGAACTGAGAGACTGGAGAGATCTAAAGTGC 942
QY 962 AACGTTTTCTATGACTTTTATGATGATGAGAGAGAGAGTAAAGTAAAGCATGG 1021
DB 943 TTTTTCATCAAGACAT-----GGAAGAGCATAT 971
QY 1022 TGAAGATCAATATTTTATATATCAACATCTTATATTTATGATTTAGTTGATPACA 1081
DB 972 GGAAGACACACACAGATTTGTTACACCCAGCTGTGTTCTTGATTCATTTGATTTCTCA 1031
QY 1082 GTGGTTAGTGAATGATTTTTCATTTCTTTCCTTGAAGTCTTACCTTCAAGTAAACA 1141
DB 1032 GGGGTATCAGTGAAGATTTCTTATTTCTTTCCTTGAAGTCTTACCTTCAAGTAAACA 1084
QY 1142 AACTCTTCATCAGGCGCATGATCTATGAGACTCTTATGAGATATCTGGGTGATTTG 1201
DB 1085 ---GCTCTTTTCTGAGAACGATGACGGGGCTGTCTCAAGTCTGTCTTGAACAT----- 1137
QY 1202 ACCCCAACCATCTCTCAAAAGCATTAATATCAATATCAGCTGTATATGTTTAAATCAGC 1261
DB 1138 -CACAAGCATTTCTCAAAAACATTAATATCAAGGACATGCTGTATATGTTTCA----- 1191
QY 1262 AGAAGCATGTTTATATGTTTGTACAAAAGAGATTTATAGGTGGGAGTGAAGATATAG 1321
DB 1192 -----CTGTCCGCTGTTTTCATCTTGTATGTGAAGGGCTTGGGGTGGGATTTGA 1245
QY 1322 ACCATGATGATGATCACTTCAAGCTATTAATTAAGATCTTAAATAGGACAGAGACT 1381
DB 1246 AGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
QY 1382 GTGAAACAGACACCTTAATATATGAGTGTATGAGTGAAGCAATCTTCTGAAAACGCA 1441
DB 1306 ACAGAAATATCATCTCTGAAAATATGATGTACAGCTCAAGAAAGCAATCCCTGAAAACACA 1365
QY 1442 AACTCTTTT 1450
DB 1366 GACTCTTTT 1374
```



RESULT 6  
US-09-764-891-5477  
; Sequence 5477, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5477  
; LENGTH: 13862  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-5477

Query Match 10.5%; Score 295.2; DB 10; Length 13862;  
Best Local Similarity 74.6%; Pred. No. 5,6e-52;  
Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1667 AAACAGATGATTAAGCTGCGAAACGCTGCAACACGCTCATAGTTCTAGCTTGGAGG 1726  
DB 1508 AAACCTTAATAGCTGCTGGCTCTGTGGCTCATGCTGTGATCCAGCACTTTGGAGG 1567  
QY 1727 TTGAGAGGAGAGATGGCTTGAACAAGGTTTCAAGGCGACGCTGGGCAATTAACAG 1786  
DB 1568 CTGAGGAGAGAGATCACTTGAGGTGAGGTTTGAGACCACTGGGCAATGAGCA 1627  
QY 1787 ATCTGTCTCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1845  
DB 1628 ACCCATCTCTACTATAAAATCAAAAGTTAGCCAGGATGGGCGGCGATGGTGGG 1687  
QY 1846 TCACGCTGTATCCAGCACTTTGGAGGCGGAGCGGAGTCACTGTGTCAAGA 1905  
DB 1688 TCATGCTGTATCCAGCACTTTGGAGGCGGAGTGGGCGGATCACTGTGTCAAGA 1747  
QY 1906 GTTTGAGACCAAGCTGCGCAATGCGCAAAACCCGCTGTCTACTCAAAATGCAAAATTA 1965  
DB 1748 GTTCAAGACCAAGCTGCGCAATGCGCAAAACCCGCTGTCTACTCAAAATGCAAAATTA 1807  
QY 1966 GCCAGGCGTGTGAGCAAGCTGTATCCAGCTACTTGGAGGCTGAGGAGAGAGAT 2025  
DB 1808 GCCAGGCTTATGATGAGGCGCTGTATCCAGCTACTTGGAGGCTGAGGAGAGAT 1867  
QY 2026 CGCTTAACCCAGAGAGTGGAGTTCAGATGAGTGAATCGTCCGTTGCACTCCAGC 2085  
DB 1868 CGTTGAAGCCAGAGGCGGAGTTTGACACCGAGAGATGTGGCAGCTGCACTCCAGC 1927  
QY 2086 TGGGCGCAAGACCAAGCTCTGTCTCAAGAAAAAAGAGAGAGAGAGAGAGAGAGAG 2145  
DB 1928 TGGGCGGTAA-AGCGAGACTCTGTCTCAAGAAAAAAGAGAGAGAGAGAGAGAGAG 1986  
QY 2146 AGAGAACATATTGGAGAGAGAGATGGGAGAGCTTGCAGAAATTTGTG 2197  
DB 1987 AGAAAAAATTAGCCAGGCTGTGTCATGCACTGTGCTTACTTGGG 2038

RESULT 7  
US-09-764-891-10204  
; Sequence 10204, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10204  
; LENGTH: 13862  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-891-10204

Query Match 10.5%; Score 295.2; DB 10; Length 13862;  
Best Local Similarity 74.6%; Pred. No. 5,6e-52;  
Matches 397; Conservative 0; Mismatches 133; Indels 2; Gaps 2;

QY 1667 AAACAGATGATTAAGCTGCGAAACGCTGCAACACGCTCATAGTTCTAGCTTGGAGG 1726  
DB 1508 AAACCTTAATAGCTGCTGGCTCTGTGGCTCATGCTGTGATCCAGCACTTTGGAGG 1567  
QY 1727 TTGAGAGGAGAGATGGCTTGAACAAGGTTTCAAGGCGACGCTGGGCAATTAACAG 1786  
DB 1568 CTGAGGAGAGAGATCACTTGAGGTGAGGTTTGAGACCACTGGGCAATGAGCA 1627  
QY 1787 ATCTGTCTCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1845  
DB 1628 ACCCATCTCTACTATAAAATCAAAAGTTAGCCAGGATGGGCGGCGATGGTGGG 1687  
QY 1846 TCACGCTGTATCCAGCACTTTGGAGGCGGAGCGGAGTCACTGTGTCAAGA 1905  
DB 1688 TCATGCTGTATCCAGCACTTTGGAGGCGGAGTGGGCGGATCACTGTGTCAAGA 1747  
QY 1906 GTTTGAGACCAAGCTGCGCAATGCGCAAAACCCGCTGTCTACTCAAAATGCAAAATTA 1965  
DB 1748 GTTCAAGACCAAGCTGCGCAATGCGCAAAACCCGCTGTCTACTCAAAATGCAAAATTA 1807  
QY 1966 GCCAGGCGTGTGAGCAAGCTGTATCCAGCTACTTGGAGGCTGAGGAGAGAGAT 2025  
DB 1808 GCCAGGCTTATGATGAGGCGCTGTATCCAGCTACTTGGAGGCTGAGGAGAGAT 1867  
QY 2026 CGCTTAACCCAGAGAGTGGAGTTCAGATGAGTGAATCGTCCGTTGCACTCCAGC 2085  
DB 1868 CGTTGAAGCCAGAGGCGGAGTTTGACACCGAGAGATGTGGCAGCTGCACTCCAGC 1927  
QY 2086 TGGGCGCAAGACCAAGCTCTGTCTCAAGAAAAAAGAGAGAGAGAGAGAGAGAGAG 2145  
DB 1928 TGGGCGGTAA-AGCGAAGCTGTCTCAAGAAAAAAGAGAGAGAGAGAGAGAGAG 1986  
QY 2146 AGAGAACATATTGGAGAGAGAGATGGGAGAGCTTGCAGAAATTTGTG 2197  
DB 1987 AGAAAAAATTAGCCAGGCTGTGTCATGCACTGTGCTTACTTGGG 2038

RESULT 8  
US-10-205-428-1003  
; Sequence 1003, Application US/10205428  
; Publication No. US20030108907A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P117C1  
; CURRENT APPLICATION NUMBER: US/10/205,428  
; CURRENT FILING DATE: 2002-07-26  
; PRIOR APPLICATION NUMBER: 09/764,892  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11





Db 43446 AGCGGAGGTTGAGAGGACCAAGATCGGCGACTGTATCCAGCTGGGTGACA-AGC 43504  
QY 2100 AAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGA 2142  
Db 43505 AAGACTCCATCTCAAAAAAAGAGAGAGAGAGA 43547

## RESULT 13

US-09-954-531-180  
; Sequence 180, Application US/09954531  
; Patent No. US20020165180A1  
; GENERAL INFORMATION:  
; APPLICANT: Weaver, Zoe  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
; FILE REFERENCE: 689290-77  
; CURRENT APPLICATION NUMBER: US/09/954,531  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: US/60/233,133  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,009  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,034  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,509  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: US/60/234,567  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 1392  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 180  
; LENGTH: 65608  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(65608)  
; OTHER INFORMATION: n=a,l,g or c  
US-09-954-531-180

Query Match 10.0%; Score 281.4; DB 9; Length 65608;  
Best Local Similarity 81.2%; Pred. No. 1.2e-48;  
Matches 376; Conservative 0; Mismatches 81; Indels 6; Gaps 4;

QY 1682 GCTGGAACCGTGGACAGCTCATAGTCTGTAGCTGTTGGAGGTTGAGAGGAGAT 1741  
Db 43089 GCGGGGTGTGTGCTCAGCTGTATCCAGCACTTGGAGTCTGAGGAGAGAT 43148  
QY 1742 GCGTTAACAAGTGTTCAGGCGCAGCTGGCAACATACAGATCCTGTCTCTCAA 1801  
Db 43149 CCTTGAAGCCAGAAATTCA-GCCAGCTTGGCAATGTAGGAGACGCATCTTACTA 43207  
QY 1802 AAAAAAAAAAAAAAGA-AAAGAGAGGCGCGGCGTGTGTGCTCAAGCTGTATC 1859  
Db 43208 AAAAAAAAAAAAAAGAAGAAATTTAGCGCGGCGTGTGTGCTCAAGCTGTATC 43267  
QY 1860 CCAGCACTTTGGAGGCCAGCGCGGCGATCACCTGTGTGAGAGTTGAGACCAAGC 1919  
Db 43268 CCAGCACTTTGGAGGCCAGCGCGGCGATCAC-GAGGTGAGAGTTGAGACCAAGC 43325  
QY 1920 TGGCCAACTGCAAAACCCGCTGTACTCAAAATGCAAAATTAGCCAGGCGTGTAG 1979  
Db 43326 TGGCCAACTGCAAAACCCGCTGTACTCAAAATGCAAAATTAGCCAGGCGTGTAG 43385  
QY 1980 CAGGCACTGTATCCAGCTACTTGTGAGGCTGAGGAGAAATCGCTTGAACCAAG 2039  
Db 43386 CCGGCGCTGTATCCAGCTACTTGTGAGGCTGAGGAGAAATCGCTTGAACCAAG 43445  
QY 2040 AGGTGAGGTTGAGTGAAGTGAATGTCGCGCTTGCATCCAGCTTGGGCAAGAGC 2099  
Db 43446 AGCGGAGGTTGAGTGAAGTGAATGTCGCGCTTGCATCCAGCTTGGGCAAGAGC-AGC 43504

QY 2100 AAGACTCTGTCTCAGAAAAAAGAGAGAGAGAGA 2142  
Db 43505 AAGACTCCATCTCAAAAAAAGAGAGAGAGAGA 43547

## RESULT 14

US-09-764-877-2857  
; Sequence 2857, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2857  
; LENGTH: 11821  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-2857

Query Match 9.9%; Score 280.2; DB 9; Length 11821;  
Best Local Similarity 74.4%; Pred. No. 8e-49;  
Matches 386; Conservative 0; Mismatches 118; Indels 15; Gaps 2;

QY 1635 AGTATTTACTATATTAAGAGGTTGTGCAAAACAGATGATTAAGTGGCAACGCTG 1694  
Db 4599 AGGAACTGCAAACTATGTTCAGCAGCAAGAAATGAAATGAAAGCCAGCATGATG 4658  
QY 1695 GCAACGCTCATAGTCTTCTAGCTGTTGAGAGTTGAGAGAGAGATGCTTGAACAG 1754  
Db 4659 GCTATGCCGTATATCCAGCACTTTAGAG-ATTAGCAGAGAGATCACTTAAAGCCAG 4717  
QY 1755 GTGTTAAGGCGCAGCTGGCAACATACAGATCTGTCTCAAAAAAAGAGAGAGAG 1814  
Db 4718 GAGTTCAAGACCAAGTCTGGGCAATATAGTAACTTGTCTCAAAAAAAGAGAGAG 4776  
QY 1815 AAAAAAGAGAGAGGCGCGGTGTGTGTCTCAAGCTGTATCCAGATCTTGGAG 1874  
Db 4777 -----TTGCGAGGCGGTGTGTGTCTCAAGCTGTATCCAGATCTTGGAG 4823  
QY 1875 GCGAGCGCGGCGATCACTGTGTGAGAGTTGAGACCAAGCTGGCCAAATGGCAA 1934  
Db 4824 GCGAGCGGCGGATTAAGTGTGAGAGTTGAGAGTTGAGACCAAGCTGGCCAAATGGCAA 4883  
QY 1935 AAGCCGCTGTACTCAAAATGCAAAATTTAGCCAGCGTGTGAGAGCACTGTATC 1994  
Db 4884 AAGCCGCTGTACTCAAAATGCAAAATTTAGCTGGCGGTGTGAGAGCACTGTATC 4943  
QY 1995 CCAGCTACTTGGAGGCTGAGGAGAGATGCTTGAACCAAGAGTGTGAGGCTGAG 2054  
Db 4944 CCAGCTACTTGGAGGCTGAGGAGAGATCACTTGAATCTGGAGGCTGAGGAGTGTGAG 5003  
QY 2055 TAAGCTAGATCGTCCGCTGCACTCAAGCTGGGAGCAAGAGCAAGCTGTCTCAG 2114  
Db 5004 TGAAGCGGAGTGGACCATTAACCTACAGCTGGGAGCAAGAGGAGTAACTGTCTCAA 5063  
QY 2115 AAAAAAAG 2153  
Db 5064 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5102

RESULT 15  
US-10-242-515-2857  
; Sequence 2857, Application US/10242515  
; Publication No. US20040009488A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005C1

Job time : 971 secs

```

; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2857
; LENGTH: 11821
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-515-2857

```

```

Query Match          9.9%; Score 280.2; DB 15; Length 11821;
Best Local Similarity 74.4%; Pred. No. 8e-49;
Matches 386; Conservative 0; Mismatches 118; Indels 15; Gaps 2;

```

```

QY 1635 AGGATATTAAGTATTAAGAGAGTGTGACAAAGATGATTAAGCTGGAAACCGTG 1694
DB 4599 AGGAACTGCAAACTATGTTGACAGACAGAAAGAAATGAACTGAAAGCCAGCATGATG 4658
QY 1695 GCACACGCTCATGATGTTCTAGCTGTTGGAGGTTGAGAGAGAGATGCTTGAACACAG 1754
DB 4659 GCTCATGCGCGTATATCCAGCACTTAGGA-ATTAGCAGAGAGATCACTTGAAGCCAG 4717
QY 1755 GTGTTCAAGGCGCAGCTGAGCAACATACAGATCTGTCTCAAAAAAAAAAAAAA 1814
DB 4718 GAGTCAAGACCAAGTCTGGCAATATAGTAAGCTGTCTCTACAAAAAATGAAAAA- 4776
QY 1815 AAAAGAAAGAGAGGCGCGGCGTGTGCTCAAGCTGTATCCAGCACTTTGGAG 1874
DB 4777 -----TTGGCCAGGCGTGTGCTCAAGCTGTATTCAGCACTTTGGAG 4823
QY 1875 GCCGAGCGGCGCGATCACTGTGTCAAGAGTTGAGCCAGCTGGCCAAATGGCAA 1934
DB 4824 GCCGAGCGGCGAGATTAAGCTGTAGGTCAAGAGTTCAAGACCAAGCTGGCCAAATGGCAA 4883
QY 1935 AACCCGCTGTATCTCAAAATGCAAAATTTAGCCAGCGTGTAGCAGGCACTGTATC 1994
DB 4884 AACCCGCGCTCTAATAAATAAATAAATTAGCTGGCGGTGGGAGCACTGTATC 4943
QY 1995 CCAGCTACTTGGAGGCTGAGGAGAGAAATCGCTTGAACCAAGAGGTGAGGTTGCAG 2054
DB 4944 CCAGCTACTTGGAGGCTGAGGAGAGAAATCACTTGAATCGGAGGTGAGGTTGCAG 5003
QY 2055 TAAGCTGAGATCGTGGCGTTGCACTCAAGCTGGGCGAGAGAGCAAGCACTGTCTCAG 2114
DB 5004 TGAGCCGCGATCCACCATTAACCTACAGCTGGGCGAAGAGGTGAATCTGTCTCA 5063
QY 2115 AAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGACA 2153
DB 5064 AATTAATAATAATAATAATAATAATAATAATAATA 5102

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 06:30:42 ; Search time 7086 Seconds

(without alignment)  
11875.759 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818  
Sequence: 1 agagaaccatcaatga.....aaaaaaaaaaaaaaaa 2818

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_estch:\*  
3: em\_estcu:\*  
4: em\_estfu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	970.2	34.4	1201	13	BX402063 BX402063
2	878.8	31.2	1201	9	AL559877 AL559877
3	847	30.1	872	12	BG758510 BG758510
4	817.4	29.0	854	13	BX464578 BX464578

5	809.4	28.7	953	13	B0065440 B0065440
6	799.2	28.4	1052	13	B0555935 B0555935
7	781.2	27.7	918	13	BX391067 BX391067
8	780.8	27.7	853	13	BX464579 BX464579
9	777.8	27.6	920	13	BX392040 BX392040
10	739.4	26.2	743	12	BG686133 BG686133
11	705.6	25.0	820	12	BG757089 BG757089
12	705.2	25.0	820	12	BG755526 BG755526
13	681	24.2	633	10	BF975096 BF975096
14	667.4	23.7	942	10	BF975166 BF975166
15	663	23.5	664	12	BG754140 BG754140
16	655.4	23.3	653	12	BG757392 BG757392
17	653.2	23.2	1201	9	AL581406 AL581406
18	627.6	22.3	670	12	BG341546 BG341546
19	626	22.2	956	10	BF664355 BF664355
20	615.6	21.8	1201	13	BX402062 BX402062
21	572.6	20.3	604	10	AM978582 AM978582
22	568	20.2	928	12	BG398364 BG398364
23	562	19.9	570	10	AM504807 AM504807
24	551.4	19.6	843	12	BG758815 BG758815
25	546.8	19.4	1034	12	BG755005 BG755005
26	540	19.2	541	10	BF238155 BF238155
27	498	17.7	489	13	BX283642 BX283642
28	493.6	17.5	511	12	BG170756 BG170756
29	490.8	17.4	535	14	CD707143 CD707143
30	471	16.7	517	12	BG170824 BG170824
31	467.8	16.6	528	9	AA954956 AA954956
32	467.6	16.6	939	12	BG341819 BG341819
33	456.4	16.2	889	12	BG686876 BG686876
34	453.8	16.1	1140	10	BF664352 BF664352
35	446.2	15.8	476	10	AM135547 AM135547
36	442	15.7	442	9	A1016902 A1016902
37	441.4	15.6	464	10	AM452648 AM452648
38	412.8	14.6	1292	11	AK080144 AK080144
39	411.4	14.6	757	9	AJ446140 AJ446140
40	409.8	14.5	696	9	AJ453647 AJ453647
41	406.2	14.4	422	10	AM263139 AM263139
42	405.4	14.4	522	12	BG144705 BG144705
43	369.4	13.1	374	10	AM401901 AM401901
44	366.2	13.0	729	9	AJ450296 AJ450296
45	365.6	13.0	688	9	AJ450317 AJ450317

ALIGNMENTS

RESULT 1  
LOCUS BX402063  
DEFINITION BX402063 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
Homo sapiens cDNA clone CSDBL012YD18 5-PRIME, mRNA sequence.  
ACCESSION BX402063  
VERSION BX402063.1  
KEYWORDS GI:30626645  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
TITLES Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6672.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSDBL012DB09QPlac1cluster=6672.r. Contact :  
Peng Liang Email: filiang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600



```

QY      80  ATGAGACGCTCTTATGAGACGGAGAGATTCTTACCAATGCAAAAATGCGCTGG 139
Db      92  ATGAGACGCTCTTATGAGACGGAGAGATTCTTACCAATGCAAAAATGCGCTGG 151
QY      140 GCTAAGGGTGGGCTGAGACCTACCTGTGCTACGTATGAGAGAGGCTGACAGTCTACA 199
Db      152 GCTAAGGGTGGGCTGAGACCTACCTGTGCTACGTATGAGAGAGGCTGACAGTCTACA 211
QY      200 TCCCTTTTCACTGACCTTGGTTATCTTCCCAATPAGAACGCGCTGCCAGTGGATTCCTC 259
Db      212 TCCCTTTTCACTGACCTTGGTTATCTTCCCAATPAGAACGCGCTGCCAGTGGATTCCTC 271
QY      260 TTCTCCGCTACATCTCCGACCTGAGACCTAGAACCTTGGCGCTGCTACCGGCTCACCTGG 319
Db      272 TTCTCCGCTACATCTCCGACCTGAGACCTAGAACCTTGGCGCTGCTACCGGCTCACCTGG 331
QY      320 TTCACTCCTGAGACCCCTGCTACGACTGTGCCGACATGTGGCGCACTTTCTGAGAGGG 379
Db      332 TTCACTCCTGAGACCCCTGCTACGACTGTGCCGACATGTGGCGCACTTTCTGAGAGGG 391
QY      380 AACCCCACTCAGTGTGAGATCTTACCGCGCGCTCTACTCTGTGAGAGCCGCAAG 439
Db      392 AACCCCACTCAGTGTGAGATCTTACCGCGCGCTCTACTCTGTGAGAGCCGCAAG 451
QY      440 GCTGACCCGAGGGGCTGGCGGGCTGACCGCGCGGGGTGCAAAATGCAATCATGACC 499
Db      452 GCTGACCCGAGGGGCTGGCGGGCTGACCGCGCGGGGTGCAAAATGCAATCATGACC 511
QY      500 TTCAAGATTAATTTTACTGCTGAGAACTATTTTGTAGAAAACATGAAAGAACTTTCAA 559
Db      512 TTCAAGATTAATTTTACTGCTGAGAACTATTTTGTAGAAAACATGAAAGAACTTTCAA 571
QY      560 GCTGAGGAGGGGTGATGAAATTCAGTGTCTCTCCAGACAGCTTCGGCGCATCTT 619
Db      572 GCTGAGGAGGGGTGATGAAATTCAGTGTCTCTCCAGACAGCTTCGGCGCATCTT 631
QY      620 TTGCCCCCTGTATGAGGTTATGATGACGAGACGATTTGTTGAGATTTGATGATG 679
Db      632 TTGCCCCCTGTATGAGGTTATGATGACGAGACGATTTGTTGAGATTTGATGATG 691
QY      680 CAACCTCCAGGAATGTCACACAGATGAATATCTCTGTGAGAGACAGTGTATTAAC 739
Db      692 CAACCTCCAGGAATGTCACACAGATGAATATCTCTGTGAGAGACAGTGTATTAAC 751
QY      740 AGCTCTTAAGTCTTCTGTTTATTTTCTTCAACCTGCTCACTTCTTGAAGTTTACAGA 799
Db      752 AGCTCTTAAGTCTTCTGTTTATTTTCTTCAACCTGCTCACTTCTTGAAGTTTACAGA 811
QY      800 AAAATATTTATATACACTCTTTAAATAAGATCTATGCTTTGAAAATAGAGAGAACACA 859
Db      812 AAAATATTTATATACACTCTTTAAATAAGATCTATGCTTTGAAAATAGAGAGAACACA 871
QY      860 GGTCTGCGCAGGAGCTGCTGCAATTTGTGCAAGTTTGAATGCAACATTTGCCCTACTG 919
Db      872 GGTCTGCGCAGGAGCTGCTGCAATTTGTGCAAGTTTGAATGCAACATTTGCCCTACTG 931
QY      920 GGAATATACAGAACTGAGAGACCTGGAGAGACTCTTAAGTGTCAAGTTTCTTGAAGATT 979
Db      932 GGAATATATATACAGAACTGAGAGACCTGGAGAGACTCTTAAGTGTCTTGAAGATT 986
QY      980 TTAGTAGATGAGAGACAGAGATGATCTTAAATAAGCATGTGTGAGAGATGATGTT 1039
Db      987 WGGWGGDTGRBARBARAGWRATCTWMAAACAATGGGAGAGTAAATTTTWTAAATCTTT 1046
QY      1040 TTAATAT 1045
Db      1047 TTTTAT 1052

```

```

RESULT 3
Bg758510
LOCUS      Bg758510      872 bp      mRNA      linear      EST 15-MAY-2001

```

```

DEFINITION 602712721F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853069 5',
            mRNA sequence.
ACCESSION  BG758510
VERSION    BG758510.1 GI:14069163
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 872)
AUTHORS   NIH-MGC http://mgs.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: egsabds-remail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICM698 row: 1 column: 06
            High quality sequence stop: 836.
            Location/Qualifiers
                1..872
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4853069"
                /tissue_type="primary B-cells from tonsils (cell line)"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 48"
                /note="Organ: B-cells; Vector: pOT87; Site 1: XhoI;
                Site 2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
                for average insert size 1.8kb. Library constructed by Ling
                Hong in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: This is a NIH_MGC Library."

ORIGIN
Query Match      30.1%; Score 847; DB 12; Length 872;
Best Local Similarity 99.3%; Pred. No. 2e-100; 5; Indels 1; Gaps 1;
Matches 861; Conservative 0; Mismatches 5;

QY      16  ATGAGTAGATTTTCTGGCTGAGACTTGAGAGGAGCAAGACCTCTGACAC 75
Db      2  ATGAGTAGATTTTCTGGCTGAGACTTGAGAGGAGCAAGACCTCTGACAC 61
QY      76  CACTATGACAGCTCTTGATGACCGGAGAAATTCTTTACAAATGCAAAATGTCGG 135
Db      62  CACTATGACAGCTCTTGATGACCGGAGAAATTCTTTACAAATGCAAAATGTCGG 121
QY      136  CTGGGCTAAGGGTGGGCGGAGACCTACCTGTGCTACGTAGAGAGAGGCGTACAGTGC 195
Db      122  CTGGGCTAAGGGTGGGCGGAGACCTACCTGTGCTACGTAGAGAGAGGCGTACAGTGC 181
QY      196  TACATCCTTTTCACTGACCTTGGTTATCTTGCATATAGAAAGCGTGCACGATGAAAT 255
Db      182  TACATCCTTTTCACTGACCTTGGTTATCTTGCATATAGAAAGCGTGCACGATGAAAT 241
QY      256  GCTCTTCTCCGCTACATCTCGAGCTGGAGACCTGACCTGCGCGCTGCTACCGCTCAC 315
Db      242  GCTCTTCTCCGCTACATCTCGAGCTGGAGACCTGACCTGCGCGCTGCTACCGCTCAC 301
QY      316  CTGCTTCACTCCTGAGAGCCCTGCTACAGACTGTGCCAGCATGTGGCCGACCTTTTGGG 375
Db      302  CTGCTTCACTCCTGAGAGCCCTGCTACAGACTGTGCCAGCATGTGGCCGACCTTTTGGG 361
QY      376  AGGAGACCCCAACCTCAGTCTGAGATCTTCAACCGGCGCTCTACTTCTGTAGAGACCG 435

```



Db 362 AGGAGACCCCACTAGTCTAGATCTTCAACCGCGCCCTCTACTTCTGTGAGACCG 421  
QY 436 CAAAGCTGAGCCCGAGGCGCTGCGCGCTGCACCGCGCGCGGCGCAATATGCCATCAT 495  
Db 422 CAAAGCTGAGCCCGAGGCGCTGCGCGCTGCACCGCGCGCGGCGCAATATGCCATCAT 481  
QY 496 GACCTTCAAGATATTTTACTGCTGGAATACTTTTGTAGAAACCATGAAAGACTTT 555  
Db 482 GACCTTCAAGATATTTTACTGCTGGAATACTTTTGTAGAAACCATGAAAGACTTT 541  
QY 556 CAAAGCTGAGGAGGCGCTGATGAAATCAAGTGTCTCTCAAGAGCTTCCGCGCAT 615  
Db 542 CAAAGCTGAGGAGGCGCTGATGAAATCAAGTGTCTCTCAAGAGCTTCCGCGCAT 601  
QY 616 CTTTGGCCCTGATGAGGTTGATGACTTACAGAGAGCATTTCTACTTTGGACTTTG 675  
Db 602 CTTTGGCCCTGATGAGGTTGATGACTTACAGAGAGCATTTCTACTTTGGACTTTG 661  
QY 676 ATAGCAACTTCCAGGAATGTCACACACGATGAAATCTCTGCTGAGACAGTGAATA 735  
Db 662 ATAGCAACTTCCAGGAATGTCACACACGATGAAATCTCTGCTGAGACAGTGAATA 721  
QY 736 AAGAGCTTCAAGTCTCTCTGTTTATTTCTTCACTCTCACTTTCTTGAAGTTAC 795  
Db 722 AAGAGCTTCAAGTCTCTCTGTTTATTTCTTCACTCTCACTTTCTTGAAGTTAC 781  
QY 796 AGAAATATTTATATATAGACATCTTAAAGATATATGCTTGAATATGAGAGAA 855  
Db 782 AGAAATATTTATATATAGACATCTTAAAGATATATGCTTGAATATGAGAGAA 841  
QY 856 CACAGCTTGGCCAGGAGCGTGTGCA 882  
Db 842 CCGCGCTGG-CAGGAGCGTGTGCA 867

RESULT 4  
BX464578/c 854 bp mRNA linear EST 22-MAY-2003  
LOCUS BX464578 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA  
DEFINITION clone CS0D003YB14 3-PRIME, mRNA sequence.  
ACCESSION BX464578  
VERSION BX464578.1 GI:31029653  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 854)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by life technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 6672.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS1D001ZF10NP1&cluster=6672.r. Contact :  
Peng Liang Email: filiang@life.techn.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paradey Avenue Genoscope sequence ID : CS1D001ZF10NP1.  
FEATURES  
Source Location/Qualifiers  
1..854  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D003YB14"  
/tissue\_type="B CELLS (RAMOS CELL LINE)"  
/cell\_line="RAMOS CELL LINE"  
/clone\_id="Homo sapiens B CELLS (RAMOS CELL LINE)"

/note="Vector: PCWSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the PCWSPORT 6 vector.  
Library was not normalized."

Query Match 29.0%; Score 817.4; DB 13; Length 854;  
Best Local Similarity 96.1%; Pred. No. 1,3e-96;  
Matches 816; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 929 GAACTGAGAGACCGGAGAGATCTTAAGTCAAGCTTTTCTATGACTTTAGGAGG 988  
Db 854 GAACTGAGAGACCGGAGAGATCTTAAGTCAAGCTTTTCTATGACTTTAGGAGG 795  
QY 989 ATAGAGCAGAGGTAGATCTTAAGATGAGTGAAGAGATCAATGTTTTATCA 1048  
Db 794 ATAGAGCAGAGGTAGATCTTAAGATGAGTGAAGAGATCAATGTTTTATCA 725  
QY 1049 CATCTTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1108  
Db 734 CATCTTTATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 675  
QY 1109 CTTTCCCTGAGCTTACTTCAAGTACACAACTCTTCAAGGCGCATGATGATA 1168  
Db 674 CTTTCCCTGAGCTTACTTCAAGTACACAACTCTTCAAGGCGCATGATGATA 615  
QY 1169 GAACTCTATGAGATCTGAGTATGATTTGATTTGATTTGATTTGATTTGATTT 1228  
Db 614 GAACTCTATGAGATCTGAGTATGATTTGATTTGATTTGATTTGATTTGATTT 555  
QY 1229 ATATCCATATGAGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1288  
Db 554 ATATCCATATGAGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 495  
QY 1289 AGAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1348  
Db 494 AGAAGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 435  
QY 1349 TTAATTAAGATCTTAATTAAGTGAAGAGATCTGTAACAAGACCTTATATGAGTT 1408  
Db 434 TTAATTAAGATCTTAATTAAGTGAAGAGATCTGTAACAAGACCTTATATGAGTT 375  
QY 1409 GATGCTGAAGTACCAATCTTCTGAAACGCAACTTTTAAAGAGTCCCTAATTTA 1468  
Db 374 GATGCTGAAGTACCAATCTTCTGAAACGCAACTTTTAAAGAGTCCCTAATTTA 315  
QY 1469 GAAACCCCAAACTTCAATATCATATTTAGCAAACTTTGAGAGAGAGTCTTGAA 1528  
Db 314 GAAACCCCAAACTTCAATATCATATTTAGCAAACTTTGAGAGAGAGTCTTGAA 255  
QY 1529 TGTGGGGAGAGAAATCTATTGGCTCTGGTGGGCTCTTCAATCAGAAATGCCAATC 1588  
Db 254 TGTGGGGAGAGAAATCTATTGGCTCTGGTGGGCTCTTCAATCAGAAATGCCAATC 195  
QY 1589 AGCTCAAGTTTGTATATTTTGTATGTTGATGTTGATGTTGATGTTGATGTTGAT 1648  
Db 194 AGCTCAAGTTTGTATATTTTGTATGTTGATGTTGATGTTGATGTTGATGTTGAT 135  
QY 1649 TATTAAGAGTTGTGAACAAACAGATGATTAAGCTGCAACGCTGCAACGCTATAG 1708  
Db 134 TATTAAGAGTTGTGAACAAACAGATGATTAAGCTGCAACGCTGCAACGCTATAG 75  
QY 1709 TTTAGCTGCTTGGAGTTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1768  
Db 74 TTTAGCTGCTTGGAGTTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15  
QY 1769 CTTGGGCAA 1777  
Db 14 CTTGGGCAAM 6

RESULT 5

LOCUS	BO065440	953 bp	mRNA	linear	EST 02-APR-2007
DEFINITION	AGENCOURT.6855061 NIH_MGC_99 Homo sapiens CDNA clone IMAGE:5929977 5', mRNA sequence.				
ACCESSION	BO065440				
VERSION	BO065440.1	GI:19894486			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 953)				
TITLE	NIH-MGC http://mgs.nci.nih.gov/				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Lou Straud CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Plate: LCM2108 row: p column: 10 High quality sequence scop: 634. Location/Qualifiers				
FEATURES	source	1..953			
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:5929977"				
	/tissue_type="lymphoma, cell line"				
	/lab_host="DH10B (phage-resistant)"				
	/clone_idb="NIH MGC 99"				
	/note="Organ: lymph; Vector: pOT8; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGCAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."				
ORIGIN					
Query Match	28.7%; Score 809.4; DB 13; Length 953;				
Best Local Similarity	98.5%; Pred. No. 1.3e-95;				
Matches	828; Conservative 0; Mismatches 11; Indels 2; Gaps 1;				
QY	17 TTGAAGTGAATTTTCTGAGCTGAGACTTGAGAGGAGCAAGACACTCTGACACC	76			
DB	1 TTGAAGTGAATTTTCTGAGCTGAGACTTGAGAGGAGCAAGACACTCTGACACC	60			
QY	77 ACTATGACACGCTTTGATGAAACCGAGGAAGTTCTTTACCAATTCAAAATGTCGC	136			
DB	61 ACTATGACACGCTTTGATGAAACCGAGGAAGTTCTTTACCAATTCAAAATGTCGC	120			
QY	137 TGGGTTAGGGGCGCGTGAAGACCTGAGCTACGATGAGGAAGGCGTGAACAGTCT	196			
DB	121 TGGGTTAGGGTGGCGGTGAAGACCTGAGCTACGATGAGGAAGGCGTGAACAGTCT	180			
QY	197 ACATCCCTTTGACAGCTTTGTTATCTTGCAATTAAGACGCGTGCACGTGAATG	256			
DB	181 ACATCCCTTTGACAGCTTTGTTATCTTGCAATTAAGACGCGTGCACGTGAATG	240			
QY	257 CTCTTCTTCGCTACATTCGGAAGTGGGACCTGACCTGACCGTCTACCGGCTACC	316			
DB	241 CTCTTCTTCGCTACATTCGGAAGTGGGACCTGACCTGACCGTCTACCGGCTACC	300			
QY	317 TGATTACCTCTGAGAGCGCGTGAAGACCTGTCGCAATGATGAGCGCACTTTCTGGG	376			
DB	301 TGATTACCTCTGAGAGCGCGTGAAGACCTGTCGCAATGATGAGCGCACTTTCTGGG	360			

FEATURES	source
RESULT 6	
LOCUS	BQ055935
DEFINITION	BQ055935 1052 bp mRNA linear EST 29-MAR-2002
ACCESSION	AGNCOURT_6796291 NIH_MGC_99 Homo sapiens CDNA clone IMAGE:5808181
VERSION	BQ055935
KEYWORDS	BQ055935.1 GI:19815262
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 1052)
COMMENT	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Straube, Ph.D. Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a> Tissue Procurement: Lou Straub CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LCM2051 row: m column: 14 High quality sequence, stop: 665. Location/Qualifiers
	1..1052
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:5808181"
	/issue_type="lymphoma, cell line"
	/lab_host="DH10B (phage-resistant)"

/note=lib="NIH\_MGC\_99"  
 EcorRI, cDNA made by oligo-dT priming. Directionally cloned  
 into EcorRI/XhoI sites using the following 5' adaptor:  
 GGACGAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using Zap-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."

## ORIGIN

Query Match 28.4%; Score 799.2; DB 13; Length 1052;  
 Best Local Similarity 99.2%; Pred. No. 2.6e-94;  
 Matches 823; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

17 TTGAAGTGAATTTTCTGGCTGAGACTTGAGGAGGAGGAGAGACCTCTGAGACCC 76  
 1 TTGAAGTGAATTTTCTGGCTGAGACTTGAGGAGGAGGAGAGACCTCTGAGACCC 60  
 77 ACTATGAGACGCTCTTGATGAACCGAGAGAACTTTCTTACCAATTCAAAATGTCGC 136  
 61 ACTATGAGACGCTCTTGATGAACCGAGAGAACTTTCTTACCAATTCAAAATGTCGC 120  
 137 TGGGCTAAGGCTGGGCTGAGACCTTCTGTCTACGTAGTAAGAGGCGGTGACAGTCT 196  
 121 TGGGCTAAGGCTGGGCTGAGACCTTCTGTCTACGTAGTAAGAGGCGGTGACAGTCT 180  
 197 ACATCCTTTTCACTGACCTTGGTTATCTTCGCAATAAAGCGCTGCCACCTGGAATTG 256  
 181 ACATCCTTTTCACTGACCTTGGTTATCTTCGCAATAAAGCGCTGCCACCTGGAATTG 240  
 257 CTCTTCTCCGCTACATCTCGACTGAGACCTAGACCTGAGCGGTGCTACCGCTCAC 316  
 241 CTCTTCTCCGCTACATCTCGACTGAGACCTAGACCTGAGCGGTGCTACCGCTCAC 300  
 317 TGGTTACCTCTGAGAGCCCTGCTACGACTGCTGCCGACATGTGGCGACTTTCTGCGA 376  
 301 TGGTTACCTCTGAGAGCCCTGCTACGACTGCTGCCGACATGTGGCGACTTTCTGCGA 360  
 377 GGGAAACCCCACTCACTGAGAGATCTTCAACGCGCGCTCTTACTTCTGTAGAGACCGC 436  
 361 GGGAAACCCCACTCACTGAGAGATCTTCAACGCGCGCTCTTACTTCTGTAGAGACCGC 420  
 437 AAGCTGAGGCGGAGGAGGCTGCGGCGGCTGACCGCGCGGAGGAGTAAGCATCAG 496  
 421 AAGCTGAGGCGGAGGAGGCTGCGGCGGCTGACCGCGCGGAGGAGTAAGCATCAG 480  
 497 ACCTCAAGATTAATTTTACTGCTGAAATACCTTTGTAGAAACCATGAAAGACTTTC 556  
 481 ACCTCAAGATTAATTTTACTGCTGAAATACCTTTGTAGAAACCATGAAAGACTTTC 540  
 557 AAGGCTGGGAGAGGCTGCTGATAAATTCAATTGCTTCTGACAGAGCTTGGCGCATC 616  
 541 AAGGCTGGGAGAGGCTGCTGATAAATTCAATTGCTTCTGACAGAGCTTGGCGCATC 600  
 617 CTTTTCCCTGATATAGGTTGATGACTTACAGAGAGCATTTGATTTGGGACTTTGA 676  
 601 CTTTTCCCTGATATAGGTTGATGACTTACAGAGAGCATTTGATTTGGGACTTTGA 660  
 677 TAGCACTTCAAGAGATGTCACACAGATGAATAATCTGCTGTAAGACAGGATTAANA 736  
 661 TAGCACTTCAAGAGATGTCACACAGATGAATAATCTGCTGTAAGACAGGATTAANA 720  
 737 AACAGTCC-TTCAAGTCTTCTGTTTATTTTCAACTCTCACTTTCTTAGAGTTTAC 795  
 721 AACAGTCTTCAAGCTTCTGTTTATTTTCAACTCTCACTTTCTTAGAGTTTAC 780  
 796 AGAAGAAATATTTATATAGACTTTT-AAAGAGCATATGCTGAGAAA 844  
 781 AGAAGAAATATTTATATAGACTTTTAAAGAGCATATATGCTTGA AAA 830

RESULT 7  
 EX391067 918 bp mRNA 1linear EST 13-MAY-2003  
 LOCUS  
 DEFINITION  
 EX391067 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
 accession  
 EX391067 Homo sapiens cDNA clone CS01012YD18 5-PRIME, mRNA sequence.  
 VERSION  
 EX391067.1 GI:30615325  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eumetazoa; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 918)  
 L1, W.B., Gruber, C., Jessee, J., and Polyes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 6672.r For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0BA061ZH04\_CS05792\_1cluster=6672.r.  
 Contact: Feng Liang Email: fliang@life.techn.com URL:  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID: CS0BA061ZH04\_CS05792\_1.

## FEATURES

Source

1..918  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS01012YD18"  
 /cell\_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
 /cell\_line="RAMOS CELL LINE"  
 /clone\_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 27.7%; Score 781.2; DB 13; Length 918;  
 Best Local Similarity 95.0%; Pred. No. 6.1e-92;  
 Matches 872; Conservative 0; Mismatches 38; Indels 8; Gaps 6;

756 TCTGTTTATTTCTTCACTCTCACTCTTCTTAGAGTTTA-CAGAAAAATATTTATATA- 813  
 2 TTTTATTTTCTTACTCTCACTTTTCTTAGAGTTTACGAAAAATATTTATATA- 61  
 814 CGACTCTTAAAGATCTATGCTTGAATAATGAGAGAGACAGGCTTGGCCAGGA 873  
 62 CGACTCTTAAAGATCTATGCTTGAATAATGAGAGAGACAGGCTTGGCCAGGA 121  
 874 CGGCTGCAATGGTGCAGATTTGATGACAATGTCCTCCCTACCTGGGAATACAGAACT 933  
 122 CGGCTGCAATGGTGCAGATTTGATGACAATGTCCTCCCTACCTGGGAATACAGAACT 181  
 934 GAGAGACTGGAGACATCTTAAAGTCAAGCTTTTCTATGACTTTTATGATGAG 993  
 182 GAGAGACTGGAGACATCTTAAAGTCAAGCTTTTCTATGACTTTTATGATGAG 241  
 994 AGAGAGAGTATCTTAAAGCATGTAAGAGATCAAAATGTTTATATCAACATCC 1053  
 242 AGAGAGAGTATCTTAAAGCATGTAAGAGATCAAAATGTTTATATCAACATCC 301  
 1054 TTTATATTTGATTTTATGATTAACAGTGGGTTAGATGATATTTTCTATCTTT 1113  
 302 TTTATATTTGATTTTATGATTAACAGTGGGTTAGATGATATTTTCTATCTTT 361  
 1114 CCTTGAAGTTTACTTCAAGTATACAAACTCTTCATCAGGCCATGATCTATAGACC 1173

Db		362	CCTTGAAGTTACTTTCAAGTAACAACAACTTCATCAGGCCATGATTATAGACC	421
OY		1174	TCTTATAGAGTATCTGGGTGATTGTGACCCCAACCATCTTCCAAAGCATTAATATC	1233
Db		422	TCTTATAGAGTATCTGGGTGATTGTGACCCCAACCATCTTCCAAAGCATTAATATC	481
OY		1224	CAATATGCGCGTGATGTTTTAATCAGAGAAGCATGTTTTANOTTTGTACAAAAGAG	1293
Db		482	CAATATGCGCGTGATGTTTTAATCAGAGAAGCATGTTTTANOTTTGTACAAAAGAG	541
OY		1294	ATTGTTATGGGTGG--GATGAGGATATAGAACATGATGTCACCTTCAAGTACTTTA	1351
Db		542	ATTGTTATGGGTGGGATGGGAGGGTATAGAACATGATGTCACCTTCAAGTACTTTA	601
OY		1352	ATAAAGATCTTAAAAATGGGCAGAGAGCTGTGAACAGACACCTTAATAATGGTTGAT	1411
Db		602	ATAAAGATCTTAAAAATGGGCAGAGAGCTGTGAACAGACACCTTAATAATGGTTGAT	661
OY		1412	GCTCTAAGTAGCAATCTTCTGGGAACGCCAACTCTTTTAGAAGATCCCTAATTAGAA	1471
Db		662	GCTCTAAGTAGCAATCTTCTGGGAACGCCAACTCTTTTAGAAGATCCCTAATTAGAA	721
OY		1472	ACACCCACAACTTACATATCATATATAGCAAACATTTGG-AAGAAAGTGTCTGAATG	1539
Db		722	ACACCCACAACTTACATATCATATATAGCAAACATTTGGAAAGAAAGTGTCTGAATG	781
OY		1531	TTGGGAGAGGAAAACTAATTTGGCTCTCGTGGGTCTTTCATCTCAGAAATGCAATCAG	1590
Db		782	TTGGGAGAGGAAAACTAATTTGGCTCTCGTGGGTCTTTCATCTTGAAGAA--TGCAATCAG	839
OY		1591	GTCAGAGTTTGTCTAATTTTGTATGTGTATGCTTCTCCCAAAGTATATTAATCTATA	1658
Db		840	GTCAGAGTTTGTCTAATTTTGTATGTGTATGCTTCTCCCAAAGG-TTTATCTTTT	898
OY		1651	TAAGAGAGTTGACAAA 1668	
Db		899	TAAGAGAGTTGGACACA 916	
RESULT 8	BX464579	853 bp	mRNA	linear EST 22-MAY-2003
LOCUS	BX464579	Homo sapiens B CELLS (RAMOS CELL LINE)	Homo sapiens cDNA	
DEFINITION	BX464579	clone CS0DG003YB14.5-PRIME, mRNA sequence.		
ACCESSION	BX464579			
VERSION	BX464579.1	GI:31031641		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
JOURNAL	1 (bases 1 to 853)			
COMMENT	Li W.B., Gruber C., Jessee J. and Polayes D. Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6672.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CSIDG001ZF10QPI&cluster=6672.r. Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSIDG001ZF10QPI. Location/Qualifiers 1..853 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606"			

```

/clone="CS0DG003YB14"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_id="Homo sapiens B CELLS (RAMOS CELL LINE)"
/notes="Vector: pCMVSPORT_6, 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

Query Match	27.7%	Score 780.8	DB 13	Length 853
Best Local Similarity	96.7%	Pred. No. 7.2e-92		
Matches	801	Conservative	6	Mismatches 20
			Indels	1
			Gaps	1
QY	20	AAGTGAATTTTTCTGACCTGAGACTTGACAGAGGAGCAAGAGACACTCTGAGACCACT	79	
Db	27	AAGCAGGCTGGTACCGGTCGGAAATTCGGGATGCGAAGMGACACTCTGAGACCACT	86	
QY	80	ATGAGCAGGCTCTTGATGAACCGGAGGAAGTTTCTTACCAATTCAAAATGTCCGCTGG	139	
Db	87	ATGAGCAGGCTCTTGATGAACCGGA-GAAGTTCTTTACCAATCAAAATGTCCGCTGG	145	
QY	140	GCTAAGGATGGCGGTGAGACCTACCTGCGTACGTAGAAAGGCGGTGACAGTGCATCA	199	
Db	146	GCTAAGGATGGCGGTGAGACCTACCTGCTACGTAGTGAAGGCGGTGACAGTGCATCA	205	
QY	200	TCCCTTTCACTGACCTTGGTTATCTTGCGCAATATAGAACGGCTGCCACGTGGAATTGCTC	259	
Db	206	TCCCTTTCACTGACCTTGGTTATCTTGCGCAATATAGAACGGCTGCCACGTGGAATTGCTC	265	
QY	260	TTCCCTCGCTACATCTTGAGACTGAGACCTAGACCCTGCGCTGCTTACCGGTCACCTGG	319	
Db	266	TTCCCTCGCTACATCTTGAGACTGAGACCTAGACCCTGCGCTGCTTACCGGTCACCTGG	325	
QY	320	TTCACTCTCTGAGAGCCCTGCTGACACTGTGACCAGATGTGCGCACTTTCTGCGAGGG	379	
Db	326	TTCACTCTCTGAGAGCCCTGCTGACACTGTGACCAGATGTGCGCGCACTTTCTGCGAGGG	385	
QY	380	AAACCCCACTCAATCTGAGAGATTTCACCGCGCGCTCTAATTCTGTGAGAACCGCAAG	439	
Db	386	AAACCCCACTCAATCTGAGAGATTTCACCGCGCGCTCTAATTCTGTGAGAACCGCAAG	445	
QY	440	GCTGAGCCGAGGGGGCTGCGCGCGCTGACACGCGCGGGGTCAAATAGCATCATGACC	499	
Db	446	GCTGAGCCGAGGGGGCTGCGCGCGCTGACACGCGCGGGGTCAAATAGCATCATGACC	505	
QY	500	TTCAAAATATTTTAACTGCTGAGAACTTTGTAGAAAAACATGAAAGAACTTTCAA	559	
Db	506	TTCAAAATATTTTAACTGCTGAGAACTTTGTAGAAAAACATGAAAGAACTTTCAA	565	
QY	560	GCTGAGAGGGCTGCATGAAATTCAGTTGCTCTCAGACACGCTTGGCGCATCCTT	619	
Db	566	GCTGAGAGGGCTGCATGAAATTCAGTTGCTCTCAGACACGCTTGGCGCATCCTT	625	
QY	620	TTGGCCCTGTATGAGTGTGATGACTTACGAGACGCAATTTGTAATTTTGGCACTTTGATAG	679	
Db	626	TTGGCCCTGTATGAGTGTGATGACTTACGAGACCAATTTGTAATTTTGGCACTTTGATAG	685	
QY	680	CAACTTCAGGAATGTACAACAGATGAATATCTGCGTGAAGACAGTGAATAAAAC	739	
Db	686	CAACTTCAGGAATGTACAACAGATGAATATCTGCGTGAAGACAGTGAATAAAAC	745	
QY	740	AGTCCTCAAGTCTCTCTGTTTTTATCTTCAACTCTCACTTTCTTGAAGTTTACGAA	799	
Db	746	AGTCCTCAAGTCTCTCTGTTTTTATCTTCAACTCTCACTTTCTTGAAGTTTACGAA	805	
QY	800	AAAAATATTATACGACTCTTTAAAGATCTATGCTTGAAGAAATAG 847		
Db	806	AAAAATATTATACGACTCTTTAAAGATCTATGCTTGAAGAAATAG 853		

EX392040/c  
 LOCUS EX392040 920 bp mRNA linear EST 13-MAY-2003  
 DEFINITION EX392040 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED  
 Homo sapiens cDNA clone CS0DL012YD18 3-PRIME, mRNA sequence.  
 ACCESSION EX392040  
 VERSION EX392040.1 GI:30607824  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 920)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by life technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 6672.r for  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0BA1048ZH08\_CS04576\_1cluster=6672.r.  
 Contact: Peng Liang Email: fliang@life.com URL:  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID: CS0BA1048ZH08\_CS04576\_1.  
 Location/Qualifiers  
 1..920  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DL012YD18"  
 /cell\_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"  
 /cell\_line="RAMOS CELL LINE"  
 /clone\_11b="Homo sapiens B CELLS (RAMOS CELL LINE) COT  
 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dt)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 27.6%; Score 777.8; DB 13; Length 920;  
 Best Local Similarity 95.2%; Pred. No. 1.7e-91;  
 Matches 875; Conservative 0; Mismatches 36; Indels 8; Gaps 7;

369 TTCTGCGAGGAGAAACCCAACTGCTGAGATCTTACCGCGGCTCTACTTCTGTG 428  
 |||||  
 911 TTTTGGAGAGGAAACCCAACTGAGTCTAGG-TCTTCCCGCGGCTCTA-TTTTGTG 854  
 |||||  
 429 AGGACCGCAAGGCTGAGCCGAGGGGCTGGGGCTGACGCGCCGCGGGTCAAAATG 488  
 |||||  
 853 AGGCGCGCAAGG--TGAGCCAGAGGGGCTGGCGGNTTGA-CGCGCGGGGGTCAATAG 797  
 |||||  
 489 CCATCATGACCTTCAAGATATATTTTACTGCTGAAATCTTTGTAGAAACATGAAA 548  
 |||||  
 796 CCATCATGACC-TCAAGATATATTTTACTGCTGG-ATACCTTGTGAAAAACATG-AA 740  
 |||||  
 549 GAACTTTCAAGGCTGGGAAAGGCTGATGAAATTCAGTTGCTCTTCAGACAGCTTC 608  
 |||||  
 739 GACTTTCAAGGCTGGGAAAGGCTGATGAAATTCAGTTGCTCTTCAGACAGCTTC 680  
 |||||  
 609 GGGGCAATCCTTTGGCCCTGTATGAGGTGAGACCTTAGAGACGATTTCTACTTTGG 668  
 |||||  
 679 GGGGCAATCCTTTGGCCCTGTATGAGGTGAGACCTTAGAGACGATTTCTACTTTGG 620  
 |||||  
 669 GACTTGATGACAACTTCCAGAAATGTCACACGATGAAATATCTCTGCTGAAAGAGT 728  
 |||||  
 619 GACTTGATGACAACTTCCAGAAATGTCACACGATGAAATATCTCTGCTGAAAGAGT 560  
 |||||  
 729 GGATAAAAACAGTCTTCAAGTCTTCTGTTTTTATTTCTCAACTCTCACTTTCTTAG 788  
 |||||

Db 559 GCAATAAAAACAGTCTCTCAAGTCTTCTGTTTTTATCTTCAACTCTCACTTCTTAG 500  
 Qy 789 AGTTTACAGAAAAATATTTATATACGACTCTTTAAAAAGATCTATGCTTGAATAAG 848  
 Db 499 AGTTTACAGAAAAATATTTATATACGACTCTTTAAAAAGATCTATGCTTGAATAAG 440  
 Qy 849 GAAAGAAACAGGCTGGCCAGGAGAGTGTGAAATTTGGTGAAGTTTGAATGCAATT 908  
 Db 439 GAAAGAAACAGGCTGGCCAGGAGAGTGTGAAATTTGGTGAAGTTTGAATGCAATT 380  
 Qy 909 GTCCCTTACTGGGAATTAACAGAACTGAGACCTGGAGACATCTTAAGTGTCAAGTTT 968  
 Db 379 GTCCCTTACTGGGAATTAACAGAACTGAGACCTGGAGACATCTTAAGTGTCAAGTTT 320  
 Qy 969 TTCTATGACTTTAGGTAGAGTGAAGACAGAAAGTATCTTAAAAAGATGAGTGAAG 1028  
 Db 319 TTCTATGACTTTAGGTAGAGTGAAGACAGAAAGTATCTTAAAAAGATGAGTGAAG 260  
 Qy 1029 ATCAATGTTTTTATATCAACATCTTTATTTATTTATTTATTTATTTATTTATTT 1088  
 Db 259 ATCAATGTTTTTATATCAACATCTTTATTTATTTATTTATTTATTTATTTATTT 200  
 Qy 1089 TAGGATGAGATTTTCTATTTCTTTCCCTGACGTTTACTTTCAAGTAAACAACTCTT 1148  
 Db 199 TAGGATGAGATTTTCTATTTCTTTCCCTGACGTTTACTTTCAAGTAAACAACTCTT 140  
 Qy 1149 CCATCAAGCCAGATCTATGAGACCTCTTAATGAGATCTGGGTGATGAGACCCCAA 1208  
 Db 139 CCATCAAGCCAGATCTATGAGACCTCTTAATGAGATCTGGGTGATGAGACCCCAA 80  
 Qy 1209 ACCATCTTCCAAAGATTAATATCAATGATGCGCTGATGTTTTTAATGAGCAAGA 1268  
 Db 79 ACCATCTTCCAAAGATTAATATCAATGATGCGCTGATGTTTTTAATGAGCAAGA 20  
 Qy 1269 TGTTTTATGTTGTACAA 1287  
 Db 19 GTTTTATGTTGTACAAAAA 1

RESULT 10  
 LOCUS BG686133 743 bp mRNA linear EST 01-MAY-2001  
 DEFINITION 602638412F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4766234 5',  
 mRNA sequence.  
 ACCESSION BG686133  
 VERSION BG686133.1 GI:13917530  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 743)  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: ggaas-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
 http://image.llnl.gov  
 Plate: L10K1626 row: g column: 03  
 High quality sequence stop: 740.  
 Location/Qualifiers  
 1..743  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4766234"  
 /tissue\_type="primary B-cells from tonsils (cell line)"

```

/lab host="DH10B (phage-resistant)"
/clone lib="NIH_MGC_48"
/Note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

## ORIGIN

```

Query Match      26.2%; Score 739.4; DB 12; Length 743;
Best Local Similarity 99.9%; Pred. No. 1.8e-86;
Matches 740; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAACATCATTAATTAAGTGAATTTCTGACCTGAGACTTGACAGGAGGCAAGAAGA 63
DB 2 GAACATCATTAATTAAGTGAATTTCTGACCTGAGACTTGACAGGAGGCAAGAAGA 61
QY 64 CACTGGAACACCACTATGAGACGCTCTTGATGAACCGAGAGATTCTTTACCAATT 123
DB 62 CACTGGAACACCACTATGAGACGCTCTTGATGAACCGAGAGATTCTTTACCAATT 121
QY 124 CAAAAATGTCCTGGGCTTAAGGGTGGCGGTGAACCTAAGTGTCTAAGTGAAGAG 183
DB 122 CAAAAATGTCCTGGGCTTAAGGGTGGCGGTGAACCTAAGTGTCTAAGTGAAGAG 181
QY 184 GGGTGAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTA 243
DB 182 GGGTGAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTA 241
QY 244 CCAAGTGAATTAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTAAG 303
DB 242 CCAAGTGAATTAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTAAGTGTCTAAG 301
QY 304 CTAACCGGTACCTGTTCACTCTCTGAAGCCCTGCTAACAAGTGTGTGCGCAATGTGGC 363
DB 302 CTAACCGGTACCTGTTCACTCTCTGAAGCCCTGCTAACAAGTGTGTGCGCAATGTGGC 361
QY 364 CGACTTTCTGCGAGGAGAACCCCACTCACTGATGAGATCTTACCGCGCGCTTACTT 423
DB 362 CGACTTTCTGCGAGGAGAACCCCACTCACTGATGAGATCTTACCGCGCGCTTACTT 421
QY 424 CTGTGAGACCGGAGGCTGAGCCGAGGCGCTGCGGCGCGCTGCAACCGCGCGGCTGCA 483
DB 422 CTGTGAGACCGGAGGCTGAGCCGAGGCGCTGCGGCGCGCTGCAACCGCGCGGCTGCA 481
QY 484 AATAGCATCATGACCTTCAAGATTAATTTTACTGCTGAATATCTTTGTGAAGAACCA 543
DB 482 AATAGCATCATGACCTTCAAGATTAATTTTACTGCTGAATATCTTTGTGAAGAACCA 541
QY 544 TGAAGAACTTTCAAGGCTGGAAGGGCTGATGAATAATTCAGTGTCTCTCCAGACA 603
DB 542 TGAAGAACTTTCAAGGCTGGAAGGGCTGATGAATAATTCAGTGTCTCTCCAGACA 601
QY 604 GCTTGCGGCGATCTTTTGGCCCTGATGAGTGTATGATTAACGAGCGCATTTTCGATC 663
DB 602 GCTTGCGGCGATCTTTTGGCCCTGATGAGTGTATGATTAACGAGCGCATTTTCGATC 661
QY 664 TTTGGGACTTTGATAGCAATCTTCAGAGATGTCACACAGATGAATAATCTCTGCTGAAG 723
DB 662 TTTGGGACTTTGATAGCAATCTTCAGAGATGTCACACAGATGAATAATCTCTGCTGAAG 721
QY 724 AAGTGTGATAAAAACAAGTCC 744
DB 722 AAGTGTGATAAAAACAAGTCC 742

```

```

RESULT 11
Bg757089      820 bp      mRNA      linear      EST 15-MAY-2001
LOCUS

```

```

DEFINITION      602715124F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:485517 5',
mRNA sequence.
ACCESSION      BG757089
VERSION      BG757089.1 GI:14067742
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 820)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Straubeberg, Ph.D.
Email: cga@ds-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
http://image.lnl.gov
Plate: LINC704 row: 0 column: 06
High quality sequence stop: 675.
Location/Qualifiers

```

## FEATURES

```

1..820
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:485517"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"

```

```

/Note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

## ORIGIN

```

Query Match      25.0%; Score 705.6; DB 12; Length 820;
Best Local Similarity 96.5%; Pred. No. 3.8e-82;
Matches 753; Conservative 0; Mismatches 24; Indels 3; Gaps 3;

QY 1 AGAAGCATCATTAATTAAGTGAATTTTCTGACCTGAGACTTGACAGGAGGCAAGA 60
DB 6 AGAAGCATCATTAATTAAGTGAATTTTCTGACCTGAGACTTGACAGGAGGCAAGA 55
QY 61 AGACACTCTGGAACACCACTATGAGACGCTCTTGATGAACCGAGAGAGTTCTTTACCA 120
DB 66 AGACACTCTGGAACACCACTATGAGACGCTCTTGATGAACCGAGAGAGTTCTTTACCA 125
QY 121 ATTCAAAAATGTCCTGGGCTTAAGGGTGGCGGTGAACCTAAGTGTCTAAGTGA 180
DB 126 ATTCAAAAATGTCCTGGGCTTAAGGGTGGCGGTGAACCTAAGTGTCTAAGTGA 185
QY 181 GAGCGGTGACAGTGTACATCTTTCACTGACCTTGTGATATCTTGCAATGAACGG 240
DB 186 GAGCGGTGACAGTGTACATCTTTCACTGACCTTGTGATATCTTGCAATGAACGG 245
QY 241 CTGCGACGTGGAATGCTCTCTCCGCTACATCTGGAACCTGGAACCTGACCTGCGCG 300
DB 246 CTGCGACGTGGAATGCTCTCTCCGCTACATCTGGAACCTGGAACCTGACCTGCGCG 305
QY 301 CTGCTACCGGTACCTGTTCACTCTGAGGCGCGCTGCTAAGCTGTGCGCGCATGT 360
DB 306 CTGCTACCGGTACCTGTTCACTCTGAGGCGCGCTGCTAAGCTGTGCGCGCATGT 365
QY 361 GCGCGACTTTCTGCGAGGAGAACCCCACTCACTGATGAGATCTTCAACCGCGCGCTCTA 420

```

Db 366 GCGCGACTTCTGCGAGGGAGCCCAACTCAGTCTGAGGATCTTCAACCGCGCGCTCTA 425  
QY 421 CTTCTGTGAGACCGGAGGCTGAGCCCGAGGGGCTGCGCGGCTGCAACCGCGGGGT 480  
Db 426 CTTCTGTGAGACCGGAGGCTGAGCCCGAGGGGCTGCGCGGCTGCAACCGCGGGGT 485  
QY 481 GCAATATGCAATGACCTTCAAAATATATTTTCTGCTGGAATACCTTTGTGAAAA 540  
Db 486 GCAATATGCAATGACCTTCAAAATATATTTTCTGCTGGAATACCTTTGTGAAAA 545  
QY 541 CCAATGAAAGACTTCAAGCTGGAAGGGCTGCAATGAAATTCAGTCTGCTCCAG 600  
Db 546 CCAATGAAAGACTTCAAGCTGGAAGGGCTGCAATGAAATTCAGTCTGCTCCAG 605  
QY 601 ACAAGCTTGGCGGACCTTTTGGCCCTGTATGAGGTGATGACTTACGAGACGACTTGG 660  
Db 606 ACAAGCTTGGCGGACCTTTTGGCCCTGTATGAGGTGATGACTTACGAGACGCA-TTCG 664  
QY 661 TACTTTGGGACTTTGATGACACTTCCAGAAATGTCAACA-CGATGAAATATCTTGTCT 719  
Db 665 TACTTTGGGACTTTGATGACACTTCCAGAAATGTCAACA-CGATGAAATATCTTGTCT 724  
QY 720 GAAGACAGTGGATTA-AAAACAGTCCCTCAAGCTTCTGTTTATTTCTCAACTGC 778  
Db 725 GAAGACAGTGGATTA-AAAACAGTCCCTCAAGCTTCAAGCTTCAAGACGCGCTCAAGCCC 784

RESULT 12  
Bg755526 820 bp mRNA linear EST 15-MAY-2001  
LOCUS 602716206F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4856568 5',  
DEFINITION mRNA sequence.  
ACCESSION Bg755526  
VERSION Bg755526.1 GI:14066179  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 820)  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS Unpublished (1999)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgaabs-rc@mail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LINC1707 row: k column: 01  
High quality sequence stop: 638.  
Location/Qualifiers  
1. 820  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4856568"  
/rname\_type="primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 48"  
/note="Organ: B-cells; Vector: pOTB1; Site\_1: XhoI;  
Site\_2: EcoRI; CDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: This is a NIH\_MGC Library."

ORIGIN  
Query Match 25.0%; Score 705.2; DB 12; Length 820;  
Best Local Similarity 94.5%; Pred. No. 4,3e-82;  
Matches 775; Conservative 0; Mismatches 38; Indels 7; Gaps 4;  
QY 634 GGTGATGACTTACGAGAGCATTTGGATCTTGGGACTTGTATGCAATCTCAGGAAT 693  
Db 2 GGTGATGACTTACGAGAGCATTTGGATCTTGGGACTTGTATGCAATCTCAGGAAT 60  
QY 694 GTCAACACGATGAAATATCTCTGAGACAGATGATTAATAAAGATCTTCAAGTCT 753  
Db 61 GTCAACACGATGAAATATCTCTGAGACAGATGATTAATAAAGATCTTCAAGTCT 120  
QY 754 TCTCTGTTTATTTCTTCAACTCTCACTTCTTGAAGTTTACGAAAAATTTATATA 813  
Db 121 TCTCTGTTTATTTCTTCAACTCTCACTTCTTGAAGTTTACGAAAAATTTATATA 180  
QY 814 CGACTCTTAAAAAGATCTATGCTTGAATAATGAGAAAGACAGAGTCTGGCCAGGA 873  
Db 181 CGACTCTTAAAAAGATCTATGCTTGAATAATGAGAAAGACAGAGTCTGGCCAGGA 240  
QY 874 CGTCTGCAATTTGGTGCAGTTTGAATGCAACATTTGCCCTACTGAGAAATACAGACT 933  
Db 241 CGTCTGCAATTTGGTGCAGTTTGAATGCAACATTTGCCCTACTGAGAAATACAGACT 300  
QY 934 GCAGACCTGGAGACATCTTAAAGTCAACGTTTCTATGACTTTAGTGAATGAG 993  
Db 301 GCAGACCTGGAGACATCTTAAAGTCAACGTTTCTATGACTTTAGTGAATGAG 360  
QY 994 AGCAGAAAGTATCTTAAATAAGATGTGAGAGGATCAATGTTTATATCAACATCC 1053  
Db 361 AGCAGAAAGTATCTTAAATAAGATGTGAGAGGATCAATGTTTATATCAACATCC 420  
QY 1054 TTATATTTGATTCATTTAGATTTAAGTGAAGTGTATGATGATTTTCTATCTTT 1113  
Db 421 TTATATTTGATTCATTTAGATTTAAGTGAAGTGTATGATGATTTTCTATCTTT 480  
QY 1114 CCTTGAAGTTTACTTCAAGTAAACAACTCTTCATCAGGCAATGATCTATAGACC 1173  
Db 481 CCTTGAAGTTTACTTCAAGTAAACAACTCTTCATCAGGCAATGATCTATAGACC 540  
QY 1174 TCTTAATGAGATATCTGGGTGATGTGACCCCAACATCTCTCCAAAGCATTAATATC 1233  
Db 541 TCTTAATGAGATATCTGGGTGATGTGACCCCAACATCTCTCCAAAGCATTAATATC 600  
QY 1234 CAATCATGCGCTGATATTTTATCAGCAGAAAGATTTTATGTTGTACAAAGAAAG 1293  
Db 601 CAATCATGCGCTGATATTTTATCAGCAGAAAGATTTTCCATGTCCGTACAAAGAAAG 660  
QY 1294 ATTGTATGGGTGGGAGATGAGAGAT--AGACATGATGATGACCTTCAAG--CTACT 1348  
Db 661 ATTGTATGGGTGGGAGATGAGAGATGCGCGTCTCAGACATGATGATGATGATGAT 720  
QY 1349 TTAATTAAGATCTTAATAATGGCAGAGAGACTGTGAACAAGACACCTTAATAGGTT 1408  
Db 721 TTAATTAAGATCTCACAATGGGAGCAAGAGACTGTGAACAAGAAATCTTAATAGGTT 780  
QY 1409 GATGTCGAGG-TAGCAATCTTCTGAAACGCAAACTCT 1447  
Db 781 GATGTCGAGGTTAGCAAAATCTCTGAAACGCAAACTCT 820

RESULT 13  
Bg975096 693 bp mRNA linear EST 22-JAN-2001  
LOCUS 602245679F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4336722 5',  
DEFINITION mRNA sequence.  
ACCESSION Bg975096  
VERSION Bg975096.1 GI:12342311  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens



REFERENCE  
1 (bases 1 to 693)  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LNCM209 row: n column: 19  
High quality sequence stop: 692.  
Location/Qualifiers

## FEATURES

source

1..693  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4336722"  
/issue\_type="primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_48"  
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 24.2%; Score 681; DB 10; Length 693;  
Best Local Similarity 99.9%; Pred. No. 6,5e-79;  
Matches 692; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

304 CTACCCCGTCACTGTTGTTACCTCTCTGGAGCCCTCTGCTACGCTGTGCCGACATGTGGC 363  
1 CTACCCCGTCACTGTTGTTACCTCTCTGGAGCCCTCTGCTACGCTGTGCCGACATGTGGC 60  
364 CGACTTTCTGCGAGGGAGCCCACTGCTGAGGATCTCACCGGCGCTCTACTT 423  
61 CGACTTTCTGCGAGGGAGCCCACTGCTGAGGATCTCACCGGCGCTCTACTT 120  
424 CTGTGAGGACCGGAGGCTGAGCCCGAGGGGCTGCGGCGCTGCAACCGCGCGGGGTGCA 483  
121 CTGTGAGGACCGGAGGCTGAGCCCGAGGGGCTGCGGCGCTGCAACCGCGCGGGGTGCA 180  
484 AATAGCCATCATGACCTTCAAGATTATTTTACTGCTGGAATCTTTTGTAAGAACCA 543  
181 AATAGCCATCATGACCTTCAAGATTATTTTACTGCTGGAATCTTTTGTAAGAACCA 240  
544 TGAAGAACTTCAAGCCTGGAGAGGCTGCATGAAATTCAGTTCCGTCTCCAGACA 603  
241 TGAAGAACTTCAAGCCTGGAGAGGCTGCATGAAATTCAGTTCCGTCTCCAGACA 300  
604 GCTTCGGGCAATCTTTGCCCCCTGTATGAGTTGATGACTTACGAGAGCATTTTCGATC 663  
301 GCTTCGGGCAATCTTTGCCCCCTGTATGAGTTGATGACTTACGAGAGCATTTTCGATC 360  
664 TTGGGACTTTGATGACAACTTCCAGAAATGTCACACAGAAATATCTGTGTGAAG 723  
361 TTGGGACTTTGATGACAACTTCCAGAAATGTCACACAGAAATATCTGTGTGAAG 420  
724 AAGGTGGATTAAGAAAGAGTCTTCAAGTCTTCTGTGTTTATTTCTTCAACTCTCACTT 783  
421 AAGGTGGATTAAGAAAGAGTCTTCAAGTCTTCTGTGTTTATTTCTTCAACTCTCACTT 480

QY 784 CTTAGACTTTACAGAAAATATTTATATACGACTCTTTAAAGATCTATGCTTGAA 843  
DB 481 CTTAGACTTTACAGAAAATATTTATATACGACTCTTTAAAGATCTATGCTTGAA 540  
QY 844 ATAGAGAGGACACAGCTGTGGCCAGGAGCGTGTGCAATTTGTCAGTTGAATGCA 903  
DB 541 ATAGAGAGGACACAGCTGTGGCCAGGAGCGTGTGCAATTTGTCAG- TTGAATGCA 599  
QY 904 ACATTTGCCCTTACTGGAATTAACAGAACTGACGACCTGGAGCATCTTAAGTCAA 963  
DB 600 ACATTTGCCCTTACTGGAATTAACAGAACTGACGACCTGGAGCATCTTAAGTCAA 659  
QY 964 CGTTTCTTATGACTTTTACTGATGATGAGC 996  
DB 660 CGTTTCTTATGACTTTTACTGATGATGAGC 692

## RESULT 14

BP975166 942 bp mRNA linear EST 22-JAN-2001  
LOCUS 602244657F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4335639 5',  
DEFINITION mRNA sequence.  
ACCESSION BP975166  
VERSION BP975166.1 GI:12342381  
KEYWORDS EST.  
SOURCE  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
1 (bases 1 to 942)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LNCM207 row: a column: 16  
High quality sequence stop: 707.  
Location/Qualifiers

## FEATURES

source

1..942  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4335639"  
/issue\_type="primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_48"  
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 23.7%; Score 667.4; DB 10; Length 942;  
Best Local Similarity 91.3%; Pred. No. 3e-77;  
Matches 799; Conservative 0; Mismatches 61; Indels 15; Gaps 8;

QY 4 GAACATCATTAATGAGAGATTTTCTGCGCTGAGACTTGGAGGAGGAGAAAGA 63  
DB 2 GAACATCATTAATGAGAGATTTTCTGCGCTGAGACTTGGAGGAGGAGAAAGA 61  
QY 64 CACTCTGACACACATATGACAGCTCTGATGAACCGAGGAAGTTCTTACCA- T 122



```

Db      62  CACTGTGACACACCATATGAGACGCTCTGTATGAACCGAGAGATTCTTTACCAAGT 121
Qy      123  TCAAAATATGTCCTCCCTGGGCTAAGGGTGGCGGTGAACCTTACTGTGTCTACGTAGTGAAGA 182
Db      122  TCAAAATATGTCCTCCCTGGGCTAAGGGTGGCGGTGAACCTTACTGTGTCTACGTAGTGAAGA 181
Qy      183  GGGGTGACAGTGTACATCTCTTCACTGTGAATTTGTTATCTTGTGCAATAGAAAGGCT 242
Db      182  GGGGTGACAGTGTACATCTCTTCACTGTGAATTTGTTATCTTGTGCAATAGAAAGGCT 241
Qy      243  GCCACGTGAGATGCTCTTCTCCGCTACATCTCGACCTGAGACCTAGACCTGAGCGCT 302
Db      242  GCCACGTGAGATGCTCTTCTCCGCTACATCTCGACCTGAGACCTAGACCTGAGCGCT 301
Qy      303  GCTACCGCGCTACCTGTGCTCACTCTGTGAAGCCCTGTCTACGACGTGTGCCGACATGTGG 362
Db      302  GCTACCGCGCTACCTGTGCTCACTCTGTGAAGCCCTGTCTACGACGTGTGCCGACATGTGG 361
Qy      363  CGGACTTTCTGCGAGGAGAACCCGACCTCAGTCTGAGATCTTTCACCGGCGGCTCTACT 422
Db      362  CGGACTTTCTGCGAGGAGAACCCGACCTCAGTCTGAGATCTTTCACCGGCGGCTCTACT 421
Qy      423  TCTGTGAGACCGGCAAGGCTGAGACCGGAGGCTGGCGGCTGACGCGCGCGCGGCTGAC 482
Db      422  TCTGTGAGACCGGCAAGGCTGAGACCGGAGGCTGGCGGCTGACGCGCGCGCGGCTGAC 481
Qy      483  AAATAGGCATCATGACCTTCAAGATTAATTTTACTGTGGAATCTTTTGTAGAAAAC 542
Db      482  AAATAGGCATCATGACCTTCAAGATTAATTTTACTGTGGAATCTTTTGTAGAAAAC 541
Qy      543  ATGAAGAAGCTTTCAAGAGCTGGGAGAGGCTGATGATAAATTCAGTTGCTCTCTCAAGC 602
Db      542  ATGAAGAAGCTTTCAAGAGCTGGGAGAGGCTGATGATAAATTCAGTTGCTCTCTCAAGC 601
Qy      603  AGCTTGCGGCGCATCTTTTGGCCCTGTATGA--GGTATGACTTTACGAGAGCATTTGCT 661
Db      602  AGCTTGCGGCGCATCTTTTGGCCCTGTATGAAGGCTTATGAGACTTTACGAGAGCATTTGCT 661
Qy      662  ACTTTGGGACTTTATGACAACTTCCAGAGATGTCAACAGATGATAATATCTGTCTGA 721
Db      662  ACTT--GGGACTTTATGAGAA--TTCAGAGATGTCAACA--GATGAAAATCTGTGTGA 717
Qy      722  AGACAGTGATTAATAAACAAGTCTTCAAGTCTTCTGTATTTATTTCTTCAACTCTCACT 781
Db      718  AGACAGTGATTAATAAACAAGTCTTCAAGTCTTCTGTATTTATTTCTTCAAGTCTTCA 773
Qy      782  TTCTTGAAGTTTACAAAAAATTTATATAGACTCTTTAAAAAGATCTATGTCTTGA 841
Db      774  TTCTTGAAGTTTACAAAAA--ATTTGATATTAAGTCTTTAAAGATTTATGTGTGA 830
Qy      842  AAATAGAGAGAAACAAGGCTGGCCAGGAGCT 876
Db      831  AA--TATGAGAGAAACAGGCTTTGTGACAGGAGT 863

```

```

RESULT 15
Bg754140      664 bp  mRNA  1linear  EST 15-MAY-2001
LOCUS        60270968.F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:484605 5',
DEFINITION   mRNA sequence.
ACCESSION   Bg754140
VERSION     Bg754140.1 GI:14064793
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 664)
AUTHORS    NIH-MGC http://mgs.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.

```

Email: c9abbs-r@mail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LCM1686 row: b column: 22  
 High quality sequence stop: 662.  
 Location/Qualifiers

## FEATURES

```

source
1..664
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:484605"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC 48"
/notes="Organ: B-cells; Vector: pOT7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using Zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

```

## ORIGIN

```

Query Match      23.5%; Score 663; DB 12; Length 664;
Best Local Similarity 100.0%; Pred. No. 14e-76;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      634  GGTGATGACTTACAGACAGCATTTTGTACTTTGGGACTTTGATAGCAATTCAGAGAT 693
Db      2  GGTGATGACTTACAGACAGCATTTTGTACTTTGGGACTTTGATAGCAATTCAGAGAT 61
Qy      694  GTCAACACAGATGAATATCTCTGTGAGAGACAGTGATTAATAAACAAGTCTTCAAGTCT 753
Db      62  GTCAACACAGATGAATATCTCTGTGAGAGACAGTGATTAATAAACAAGTCTTCAAGTCT 121
Qy      754  TCTGTGTTTATTTCTTCAACTCTCACTTTCTTAGAGTTTACGAAAATATTTATATA 813
Db      122  TCTGTGTTTATTTCTTCAACTCTCACTTTCTTAGAGTTTACGAAAATATTTATATA 181
Qy      814  CGACTCTTAAAGAATCTATGTCTTGAATAATGAGAGACACAGAGTCTGGCCAGGGA 873
Db      182  CGACTCTTAAAGAATCTATGTCTTGAATAATGAGAGACACAGAGTCTGGCCAGGGA 241
Qy      874  CGTGCTCAATTTGGTGTGAGTTTGAATGCAACAATGTCCCTTACTGGGATACAGAACT 933
Db      242  CGTGCTCAATTTGGTGTGAGTTTGAATGCAACAATGTCCCTTACTGGGATACAGAACT 301
Qy      934  GCAGACCTTGGAGACATCTTAAAGTCAACGTTTCTATGACTTTTAGAGTGAATGAG 993
Db      302  GCAGACCTTGGAGACATCTTAAAGTCAACGTTTCTATGACTTTTAGAGTGAATGAG 361
Qy      994  AGCAGAGTGAATCTTAAAGAAGATGTGAGAGATCAAAATTTTATATACATCC 1053
Db      362  AGCAGAGTGAATCTTAAAGAAGATGTGAGAGATCAAAATTTTATATACATCC 421
Qy      1054  TTTATTTATTTGATTCATTGATGTTAACAGTGTGTAGTATGATTTTCTATCTTTT 1113
Db      422  TTTATTTATTTGATTCATTGATGTTAACAGTGTGTAGTATGATTTTCTATCTTTT 481
Qy      1114  CCTTGAAGTTTACTTTCAAGTAAACAAACCTTTCATGAGCCATGATCTATAGACC 1173
Db      482  CCTTGAAGTTTACTTTCAAGTAAACAAACCTTTCATGAGCCATGATCTATAGACC 541
Qy      1174  TCTTAATGAGATATCTGGGTGATTTGACACCCAAACATCTCTCCAAAGCATTAATATC 1233
Db      542  TCTTAATGAGATATCTGGGTGATTTGACACCCAAACATCTCTCCAAAGCATTAATATC 601

```

Qy 1234 CAAATCAGCGGCTGTATGTTTATATCAGCAGAGCATGTTTATGTTGTACAAAGAG 1293  
Db 602 CAAATCAGCGGCTGTATGTTTATATCAGCAGAGCATGTTTATGTTGTACAAAGAG 661

Qy 1294 ATT 1296  
Db 662 ATT 664

Search completed: March 7, 2004, 11:56:02  
Job time : 7092 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 11:59:52 ; Search time 10902 Seconds

(without alignments)  
11203.510 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818  
Sequence: 1 agagaaccatcatcaattga.....aaaaaaaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 2167151695 residues

Word size : 0

Total number of hits satisfying chosen parameters: 692750

Minimum DB seq length: 0  
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database :

- 1: gb\_da:\*
- 2: gb\_hcg:\*
- 3: gb\_in:\*
- 4: gb\_cm:\*
- 5: gb\_ov:\*
- 6: gb\_pac:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_scs:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_scs:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_hcg\_hum:\*
- 31: em\_hcg\_inv:\*
- 32: em\_hcg\_other:\*
- 33: em\_hcg\_mus:\*
- 34: em\_hcg\_pln:\*
- 35: em\_hcg\_rpd:\*
- 36: em\_hcg\_mam:\*
- 37: em\_hcg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_hcg\_hum:\*
- 40: em\_hcg\_mus:\*
- 41: em\_hcg\_other:\*

*Handwritten signature*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	20	0.7	20	AR064875	AR064875 Sequence
C 2	20	0.7	20	AR080000	AR080000 Sequence
C 3	20	0.7	20	AR085926	AR085926 Sequence
C 4	20	0.7	20	AR087520	AR087520 Sequence
C 5	20	0.7	20	AR093312	AR093312 Sequence
C 6	20	0.7	20	AR118970	AR118970 Sequence
C 7	20	0.7	20	AR121692	AR121692 Sequence
C 8	20	0.7	20	AR123335	AR123335 Sequence
C 9	20	0.7	20	AR141070	AR141070 Sequence
C 10	20	0.7	20	AR154115	AR154115 Sequence
C 11	20	0.7	20	AR164658	AR164658 Sequence
C 12	20	0.7	20	E12676	E12676 Anti-HLV-1
C 13	20	0.7	20	136180	136180 Sequence 16
C 14	20	0.7	20	AR213738	AR213738 Sequence
C 15	20	0.7	20	AR222466	AR222466 Sequence
C 16	20	0.7	20	AR236083	AR236083 Sequence
C 17	20	0.7	20	AR266075	AR266075 Sequence
C 18	20	0.7	20	AR274394	AR274394 Sequence
C 19	20	0.7	20	AR305124	AR305124 Sequence
C 20	20	0.7	20	AR309228	AR309228 Sequence
C 21	20	0.7	20	AR337145	AR337145 Sequence
C 22	20	0.7	20	AR343047	AR343047 Sequence
C 23	20	0.7	20	AR344936	AR344936 Sequence
C 24	20	0.7	20	AR365970	AR365970 Sequence
C 25	20	0.7	20	AR382312	AR382312 Sequence
C 26	20	0.7	20	AR429653	AR429653 Sequence
C 27	20	0.7	20	AX004876	AX004876 Sequence
C 28	20	0.7	20	AX045779	AX045779 Sequence
C 29	20	0.7	20	AX045787	AX045787 Sequence
C 30	20	0.7	20	AX045790	AX045790 Sequence
C 31	20	0.7	20	AX104034	AX104034 Sequence
C 32	20	0.7	20	AX104364	AX104364 Sequence
C 33	20	0.7	20	AX104368	AX104368 Sequence
C 34	20	0.7	20	AX115919	AX115919 Sequence
C 35	20	0.7	20	AX133853	AX133853 Sequence
C 36	20	0.7	20	AX196224	AX196224 Sequence
C 37	20	0.7	20	AX196239	AX196239 Sequence
C 38	20	0.7	20	AX354974	AX354974 Sequence
C 39	20	0.7	20	AX355810	AX355810 Sequence
C 40	20	0.7	20	AX355811	AX355811 Sequence
C 41	20	0.7	20	AX440125	AX440125 Sequence
C 42	20	0.7	20	AX440140	AX440140 Sequence
C 43	20	0.7	20	AX465311	AX465311 Sequence
C 44	20	0.7	20	AX465326	AX465326 Sequence
C 45	20	0.7	20	AX547087	AX547087 Sequence

ALIGNMENTS

RESULT 1  
AR064875/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES

AR064875  
Sequence 5 from patent US 5849480.  
AR064875.1 GI:5995091

Unknown.  
Unclassified.  
1 (bases 1 to 20)  
Cros,P., Kufurst,R., Battall,N. and Pigs,N.  
Process and device for assaying a hapten  
Patent: US 5849480-A 5 15-DEC-1998;  
Location/Qualifiers

20 bp  
DNA  
linear  
PAT 29-SBP-1999

source 1.20  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 2  
LOCUS AR080000 20 bp DNA linear PAT 31-AUG-2000  
DEFINITION Sequence 83 from patent US 5968524.  
ACCESSION AR080000  
VERSION AR080000.1 GI:10006735  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unassigned.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Watson,J.D. and Tan,P.L.J.  
TITLE Methods and compounds for the treatment of immunologically-mediated psoriasis  
JOURNAL Patent: US 5968524-A 83 19-OCT-1999;  
FEATURES Location/Qualifiers  
source 1.20  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 3  
LOCUS AR085926 20 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 83 from patent US 5985287.  
ACCESSION AR085926  
VERSION AR085926.1 GI:10012692  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unassigned.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Tan,P., Skinner,M. and Prestidge,R.  
TITLE Compounds and methods for treatment and diagnosis of mycobacterial infections  
JOURNAL Patent: US 5985287-A 83 16-NOV-1999;  
FEATURES Location/Qualifiers  
source 1.20  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 4  
LOCUS AR087520 20 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 1 from patent US 5986084.  
ACCESSION AR087520  
VERSION AR087520.1 GI:10014283  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unassigned.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Bitch,S., Weiss,P.A. and Jenny,L.  
TITLE Ribonucleoside-derivative and method for preparing the same  
JOURNAL Patent: US 5986084-A 1 16-NOV-1999;  
FEATURES Location/Qualifiers  
source 1.20  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 5  
LOCUS AR093312 20 bp DNA linear PAT 08-SEP-2000  
DEFINITION Sequence 83 from patent US 6001361.  
ACCESSION AR093312  
VERSION AR093312.1 GI:10020062  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unassigned.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Tan,P., Hiya,J., Visser,E., Skinner,M., Scott,L. and Prestidge,R.  
TITLE Mycobacterium vaccae antigens  
JOURNAL Patent: US 6001361-A 83 14-DEC-1999;  
FEATURES Location/Qualifiers  
source 1.20  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 6  
LOCUS AR118970 20 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 96 from patent US 6150092.  
ACCESSION AR118970  
VERSION AR118970.1 GI:14100880  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unassigned.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Uchida,K., Uchida,T., Tanaka,Y., Matsuda,Y. and Kondo,S.  
TITLE Antisense nucleic acid compound targeted to VEGF  
JOURNAL Patent: US 6150092-A 96 21-NOV-2000;

```

FEATURES                               Location/Qualifiers
source                                  1..20
                                         /organism="unknown"
                                         /mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 7
LOCUS AR121692 20 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 83 from patent US 6160093.
ACCESSION AR121692
VERSION AR121692.1 GI:14105268
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
AUTHORS Visser, E.
TITLE 1 (bases 1 to 20)
JOURNAL Compounds and methods for treatment and diagnosis of mycobacterial
FEATURES Patent: US 6160093-A 83 12-DEC-2000;
source Location/Qualifiers
1..20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 8
LOCUS AR123335 20 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6169176.
ACCESSION AR123335
VERSION AR123335.1 GI:14108301
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
AUTHORS Bruce, T.C. and Dev, A.P.
TITLE Deoxynucleic alkyl thioarea compounds and uses thereof
JOURNAL Patent: US 6169176-A 1 02-JAN-2001;
FEATURES Location/Qualifiers
1..20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

```

```

RESULT 9
LOCUS AR141070/c 20 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6207819.
ACCESSION AR141070
VERSION AR141070.1 GI:14483566
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
AUTHORS Manoharan, N. and Waier, M.A.
TITLE Compounds, processes and intermediates for synthesis of mixed
JOURNAL backbone oligomeric compounds
FEATURES Patent: US 6207819-A 1 27-MAR-2001;
source Location/Qualifiers
1..20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 10
LOCUS AR154115/c 20 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 14 from patent US 6238865.
ACCESSION AR154115
VERSION AR154115.1 GI:15122168
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
AUTHORS Huang, Z. and Szostak, J.W.
TITLE Simple and efficient method to label and modify 3'-termini of RNA
JOURNAL using DNA polymerase and a synthetic template with defined overhang
FEATURES Patent: US 6238865-A 14 29-MAY-2001;
source Location/Qualifiers
1..20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 11
LOCUS AR164658 20 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 13 from patent US 6274321.
ACCESSION AR164658
VERSION AR164658.1 GI:16237754
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES Patent: US 6274321-A 1 27-OCT-2001;
source Location/Qualifiers
1..20
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match                           0.7%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

```

AUTHORS Blumberg,B.  
TITLE High throughput functional screening of cDNAs  
JOURNAL Patent: US 6274321-A 13 14 -AUG-2001;  
FEATURES Location/Qualifiers  
SOURCE 1..20  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 1 AAAAAAAAAAAAAAAAAA 20  
RESULT 12  
E12676/c 20 bp DNA linear PAT 27-APR-1998  
LOCUS E12676  
DEFINITION Anti-HTLV-1 antisense oligonucleotide.  
ACCESSION E12676  
VERSION E12676.1 GI:3251508  
KEYWORDS JP 1997052898-A/10.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Mizuguchi, M., Kurosaki, N., Makino, K., Koyanagi, Y. and Yamamoto, N.  
TITLE ANTI-HTLV-1 ANTI-SENSE OLIGONUCLEOTIDE  
JOURNAL Patent: JP 1997052898-A 10 25-FEB-1997;  
SOYAKU GIJUTSU KENKYUSHO:KK  
COMMENT OS None  
OC Artificial sequences.  
PN JP 1997052898-A/10  
PD 25-FEB-1997  
PF 09-AUG-1995 JP 1995224606  
PI MIZUGUCHI MASATOSHI, KUROSAKI NAOKO, MAKINO KEISUKE, PI  
KOYANAGI YOSHIO,  
PI YAMAMOTO NAOKI  
PC C07H21/04//A61K31/70;  
CC strandedness: Single;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: Yes;  
FH Key Location/Qualifiers  
FH source 1..20  
FT Location/Qualifiers  
FEATURES source 1..20  
/organism="Artificial sequences".  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 20 AAAAAAAAAAAAAAAAAA 1  
RESULT 13  
I36180 20 bp DNA linear PAT 13-MAY-1997  
LOCUS I36180/c  
DEFINITION Sequence 16 from patent US 5605662.  
ACCESSION I36180  
VERSION I36180.1 GI:2086693  
KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Heller, M.J. and Tu, E.  
TITLE Active programmable electronic devices for molecular biological  
JOURNAL Patent: US 5605662-A 16 25-FEB-1997;  
FEATURES Location/Qualifiers  
SOURCE 1..20  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 20 AAAAAAAAAAAAAAAAAA 1  
RESULT 14  
AR213738 20 bp DNA linear PAT 25-SEP-2002  
LOCUS AR213738  
DEFINITION Sequence 83 from patent US 6406704.  
ACCESSION AR213738  
VERSION AR213738.1 GI:23311025  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Tan, P., Visser, E., Prestidge, R. and Watson, J.D.  
TITLE Compounds and methods for treatment and diagnosis of mycobacterial  
JOURNAL Patent: US 6406704-A 93 18-JUN-2002;  
FEATURES Location/Qualifiers  
SOURCE 1..20  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 1 AAAAAAAAAAAAAAAAAA 20  
RESULT 15  
AR222466 20 bp DNA linear PAT 26-SEP-2002  
LOCUS AR222466  
DEFINITION Sequence 26 from patent US 6429300.  
ACCESSION AR222466  
VERSION AR222466.1 GI:23329997  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Kurz, M., Lohse, P. and Wagner, R.  
TITLE Peptide acceptor ligation methods  
JOURNAL Patent: US 6429300-A 26 06-AUG-2002;  
FEATURES Location/Qualifiers  
SOURCE 1..20  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. NO. 83;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1799 AAAAAAAAAAAAAAAAAA 1818  
|||  
Db 1 AAAAAAAAAAAAAAAAAA 20  
Search completed: March 7, 2004, 18:09:10  
Job time : 10905 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 11:56:07 ; Search time 1110 Seconds  
(Without alignments) 10785.065 Million cell updates/sec

Title: US-09-966-880a-7

Perfect score: 2818  
Sequence: 1 agagacacatcatatga.....aaaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1690386

Minimum DB seq length: 0  
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	20	0.7	20	AAQ25565	AAQ25565 Dye-coupled
c 2	20	0.7	20	AAQ33554	AAQ33554 Microstate
c 3	20	0.7	20	AAQ58578	AAQ58578 Sequence
c 4	20	0.7	20	AAQ94205	AAQ94205 Alpha-ano
c 5	20	0.7	20	AAQ75563	AAQ75563 Reverse t
c 6	20	0.7	20	AAQ75596	AAQ75596 Reverse t
c 7	20	0.7	20	AAQ90405	AAQ90405 T2 (synth
c 8	20	0.7	20	AAQ73704	AAQ73704 PCR prime
c 9	20	0.7	20	AAQ63649	AAQ63649 Anti-NTLV
c 10	20	0.7	20	AAV85582	AAV85582 LRPS PCR
c 11	20	0.7	20	AAV34591	AAV34591 M. vaccae
c 12	20	0.7	20	AAQ86606	AAQ86606 Oligonuc
c 13	20	0.7	20	AAQ37736	AAQ37736 Human mdm
c 14	20	0.7	20	AAQ37737	AAQ37737 Human mdm
c 15	20	0.7	20	AAQ27533	AAQ27533 Synthetic
c 16	20	0.7	20	AAQ21326	AAQ21326 Mycobacte
c 17	20	0.7	20	AAA40449	AAA40449 Electoche
c 18	20	0.7	20	AAA40448	AAA40448 Electoche
c 19	20	0.7	20	AAQ96372	AAQ96372 Primer us
c 20	20	0.7	20	AAQ35378	AAQ35378 Intersper
c 21	20	0.7	20	AAQ91117	AAQ91117 Oligonuc
c 22	20	0.7	20	AAA50193	AAA50193 2'-methox
c 23	20	0.7	20	AAQ14808	AAQ14808 Human gly

24	20	0.7	20	AAQ95176	AAQ95176 Human CDN
25	20	0.7	20	AAQ97238	AAQ97238 Phosphoro
26	20	0.7	20	AAQ87230	AAQ87230 Digoxigen
27	20	0.7	20	AAQ87241	AAQ87241 Poly T ol
28	20	0.7	20	AAQ10402	AAQ10402 DNA templ
29	20	0.7	20	AAQ16997	AAQ16997 Capture p
30	20	0.7	20	AAQ60896	AAQ60896 Conjugate
31	20	0.7	20	AAQ53428	AAQ53428 Oligonuc
32	20	0.7	20	AAQ28461	AAQ28461 Random ol
33	20	0.7	20	AAQ80891	AAQ80891 Human mdm
34	20	0.7	20	AAQ80890	AAQ80890 Human mdm
35	20	0.7	20	AAQ10371	AAQ10371 Oligonuc
36	20	0.7	20	AAQ99427	AAQ99427 Immunost
37	20	0.7	20	AAQ99099	AAQ99099 Immunost
38	20	0.7	20	AAQ99431	AAQ99431 Immunost
39	20	0.7	20	AAQ38246	AAQ38246 SNP speci
40	20	0.7	20	AAQ46465	AAQ46465 Oligonuc
41	20	0.7	20	AAQ78547	AAQ78547 Nucleotid
42	20	0.7	20	AAQ28351	AAQ28351 DNA oligo
43	20	0.7	20	AAQ01235	AAQ01235 Reverse p
44	20	0.7	20	AAQ29506	AAQ29506 Human mdm
45	20	0.7	20	AAQ29505	AAQ29505 Human mdm

#### ALIGNMENTS

AAQ25565/c	AAQ25565 standard; DNA; 20 BP.
AAQ25565	AAQ25565
25-MAR-2003 (revised)	
02-DEC-1992 (first entry)	
Dye-coupled 3'-amino modified oligonucleotide.	
DNA synthesis; RNA; antisense strands; detection; ss.	
Synthetic.	
Key	Location/Qualifiers
modified_base	20
	/*tag= a
	/note= "3-amino modified"
EP490281-A1.	
17-JUN-1992.	
06-DEC-1991; 91BP-00120935.	
11-DEC-1990; 90DE-04039488.	
(FARH ) HOECHST AG.	
Engels J, Herrlein M, Konrad R, Mag M;	
WPI; 1992-201578/25.	
New dye-coupled modified nucleosides, nucleotides and oligo:nucleotides -	
useful for synthesis of antisense DNA and RNA strands in presence of	
template, also for in-vivo and in-vitro detection of genetic material.	
Example; Page 9; 17pp; German.	
The sequence is an example of a dye coupled 3'-amino modified oligo-	
nucleotide, it can be used in the synthesis of DNA and RNA nucleosides,	
nucleotides and oligonucleotides and for the synthesis of opposite	
strands in the presence of a template strand and in fluorescence	
microscopic and macroscopic detection in vivo and in vitro of genetic	
material. It is labelled with a fluorescent dye. See also AAQ25566 and	



CC AAQ25567. (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;  
 SQ

Query Match 0.7%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
 |||||  
 DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 2  
 AAQ33554/c  
 ID AAQ33554 standard; DNA; 20 BP.

XX AAQ33554;  
 AC  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 02-FEB-1993 (first entry)

XX Microsatellite sequence from clone AGLA247.

DE PCR; selection; primers; OPTIPRM; breeding; cattle; parentage;  
 KM genetic mapping; traits; amplification; ss.

XX Bos taurus.

XX WO9213102-A1.

XX 06-AUG-1992.

XX 15-JAN-1992; 92WO-US000340.

XX 15-JAN-1991; 91US-00642342.

XX (GENM-) GENMARK.

XX Georges M, Massey JM;

XX WPI; 1992-284684/34.

PT Polymorphic bovine DNA markers - used in genetic identification, gene  
 mapping, and selective breeding.

PS Table 7; Page 150; 517pp; English.

XX The sequence is that of a bovine microsatellite sequence obtd. by  
 CC screening a library of bovine MbOI DNA fragments of between 250 and 500  
 CC bp with an (AC)15 and a (TC)15 oligonucleotide probe. One out of 50  
 CC clones cross-hybridised. Assuming independent distribution of  
 CC microsatellites and MbOI sites, the frequency of (T6)n > 9 microsatellites  
 CC in the bovine genome is estimated at >100,000. The sequence information  
 CC for ca. 230 such bovine microsatellites is summarised in the  
 CC specification and indexed herein (see below). The sequences upstream and  
 CC downstream of the microsatellite sequence were used to generate the  
 CC required PCR primers for in vitro amplification of the corresp.  
 CC microsatellite (using the program OPTIPRM). The microsatellites may be  
 CC used to identify individuals, for parentage testing, and in the genetic  
 CC mapping of economic trait loci, or genes involved in the determination of  
 CC economically important traits esp. in cattle, to allow selective  
 CC breeding. See also AAQ33501-34437. (Updated on 25-MAR-2003 to correct PN  
 CC field.)

SQ Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
 |||||

DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 3  
 AAQ58578  
 ID AAQ58578 standard; RNA; 20 BP.

XX AAQ58578;

XX 25-MAR-2003 (revised)  
 DT 21-AUG-1994 (first entry)

DE Sequence of synthetic RNA oligo which is a target nucleotide for a novel  
 DE receptor.

XX Novel receptor; nucleic acid; transport; oligo; ss.

XX Synthetic.

XX WO9404194-A1.

XX 03-MAR-1994.

XX 13-AUG-1993; 93WO-US007603.

XX 14-AUG-1992; 92US-00930087.

XX (MASI) MASSACHUSETTS INST TECHNOLOGY.

XX Usman N, Rebek J, De Mendoza J;

XX WPI; 1994-082846/10.

PT Transport of nucleic acid derivs. across membranes - using new receptors  
 PT which use salt bridging, aromatic stacking, hydrogen bonding and  
 PT chelation.

XX Example; Table 1, page 38; 103pp; English.

CC The inventors claim a method of transporting a nucleic acid deriv. across  
 CC a membrane which comprises using a receptor that uses salt bridgjn,  
 CC aromatic stacking, H bonding and chelation to recognise the nucleic acid  
 CC deriv. AAQ56305, AAQ58577-86 are nucleic acid derivs used in the  
 CC examples. (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 20 BP; 20 A; 0 C; 0 G; 0 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
 |||||  
 DB 1 AAAAAAAAAAAAAAAAAA 20

RESULT 4  
 AAQ94205/c  
 ID AAQ94205 standard; DNA; 20 BP.

XX AAQ94205;

XX 25-MAR-2003 (revised)  
 DT 24-AUG-1995 (first entry)

DE Alpha-anomeric oligonucleotide ligand 1803 for oestradiol hapten.

XX Oligonucleotide ligand; steroid hormone; hapten; immobilisation;  
 KW immunodetection; estradiol; alpha-anomer; ss.

XX Synthetic.

FX Key Location/Qualifiers

```

FT misc_feature 1. 21
FT /tag= b
FT /note= "the glycosidic bonds between nucleotides are all
FT in the alpha-anomer form"
FT modified_base
FT 20
FT /tag= a
FT /mod_base= OTHER
FT /note= "carries a group derived ffrom aminopropanediol"
XX
XX WO9429723-A1.
XX
XX 22-DEC-1994.
XX
XX 10-JUN-1994; 94WO-FR000689.
XX
XX 11-JUN-1993; 93FR-00007093.
XX
XX (CROS/) CROS P.
XX (KURF/) KURFURST R.
XX (BATT/) BATTAIL N.
XX (PIGA/) PIGA N.
XX
XX Cros P, Kurfurst R, Battail N, Piga N;
XX
XX WPI; 1995-036665/05.
XX
XX Assay device for hapten or its specific antibodies - comprises support
XX having competitive reagent immobilised via nucleic acid ligand to improve
XX orientation and accessibility.
XX
XX Example 1; Page 10; 39pp; French.
XX
XX Oligonucleotides (AAQ94201-094205) were synthesised for use as ligands.
XX The ligands are covalently linked to a hapten (esp. a steroid hormone) to
XX form a conjugate which is then immobilised on a solid support for
XX interaction with antibodies against the hapten. Nucleic acid ligands are
XX less likely to be recognised by the antibodies than are peptide ligands
XX and nucleic acids are also less likely to undergo intramolecular
XX organisation which interferes with accessibility of the hapten to the
XX antibodies. For immunodiagnosis of oestradiol, the active hapten
XX oestradiol-6-carboxymethoxime-N-hydroxysuccinimide ester was used.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;
XX
XX Query Match 0.7%; Score 20; DB 2; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 1.5e+03;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1799 AAAAAAAAAAAAAAAAAA 1818
XX DB 20 AAAAAAAAAAAAAAAAAA 1
XX
XX RESULT 5
XX AAQ75563/c
XX ID AAQ75563 standard; DNA; 20 BP.
XX
XX AAQ75563;
XX
XX 04-AUG-1995 (first entry)
XX
XX Reverse transcription primer used in cDNA analysis technique.
XX
XX Analysis; gene expression; reverse transcription; primer; cDNA;
XX aggregate; restriction enzyme; ss.
XX
XX Synthetic.
XX
XX JP06303997-A.
XX
XX 01-NOV-1994.
XX

```

```

PF 16-APR-1993; 93JP-00112515.
XX
XX 16-APR-1993; 93JP-00112515.
XX
XX (NITE ) NIPPON TELEGRAPH & TELEPHONE CORP.
XX
XX WPI; 1995-018287/03.
XX
XX Analysis of cDNA and gene expression - by amplification of mRNA followed
XX by digestion with restriction enzymes.
XX
XX Disclosure; Page 5; 11pp; Japanese.
XX
XX A method for the analysis of cDNA comprises (a) preparing an aggregate of
XX double-stranded cDNAs by using an aggregate of mRNAs and a plural type of
XX labelled reverse transcription primers (GENBSEQ files AAQ75547-075798)
XX and using the aggregate of mRNAs as the template for each reverse
XX transcription primer; (b) digesting each of the prepared aggregates of
XX the double-stranded cDNAs with restriction enzyme and; (c)
XX electrophoresing the digested aggregate of cDNAs in separate lanes. The
XX method can be used to analyse gene expression rapidly and easily
XX
XX Sequence 20 BP; 1 A; 0 C; 2 G; 17 T; 0 U; 0 Other;
XX
XX Query Match 0.7%; Score 20; DB 2; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 1.5e+03;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1796 CTCAAAAAAAAAAAAAAAAA 1815
XX DB 20 CTCAAAAAAAAAAAAAAAAA 1
XX
XX RESULT 6
XX AAQ75596/c
XX ID AAQ75596 standard; DNA; 20 BP.
XX
XX AAQ75596;
XX
XX 04-AUG-1995 (first entry)
XX
XX Reverse transcription primer used in cDNA analysis technique.
XX
XX Analysis; gene expression; reverse transcription; primer; cDNA;
XX aggregate; restriction enzyme; ss.
XX
XX Synthetic.
XX
XX JP06303997-A.
XX
XX 01-NOV-1994.
XX
XX 16-APR-1993; 93JP-00112515.
XX
XX 16-APR-1993; 93JP-00112515.
XX
XX (NITE ) NIPPON TELEGRAPH & TELEPHONE CORP.
XX
XX WPI; 1995-018287/03.
XX
XX Analysis of cDNA and gene expression - by amplification of mRNA followed
XX by digestion with restriction enzymes.
XX
XX Disclosure; Page 5; 11pp; Japanese.
XX
XX A method for the analysis of cDNA comprises (a) preparing an aggregate of
XX double-stranded cDNAs by using an aggregate of mRNAs and a plural type of
XX labelled reverse transcription primers (GENBSEQ files AAQ75547-075798)
XX and using the aggregate of mRNAs as the template for each reverse
XX transcription primer; (b) digesting each of the prepared aggregates of
XX the double-stranded cDNAs with restriction enzyme and; (c)
XX electrophoresing the digested aggregate of cDNAs in separate lanes. The
XX method can be used to analyse gene expression rapidly and easily
XX

```

XX Sequence 20 BP; 2 A; 1 C; 0 G; 17 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2792 TTGAAAAAAAAAAAAAAAAA 2811  
DB 20 TTGAAAAAAAAAAAAAAAAA 1

RESULT 7  
AAQ90405/C  
ID AAQ90405 standard; DNA; 20 BP.

AC AAQ90405;  
XX  
DT 08-JAN-1996 (first entry)

XX T2 (synthetic DNA probe with 5' amino terminal #4).

XX T2; HLA; dQa; self-addressable electronic device; SAED; hybridisation;  
XX ss.

OS Synthetic.

XX Key Location/Qualifiers  
FH misc\_feature 1

FT /tag= a  
FT /note= "3' aminolink2 Thymine; allows binding to any  
FT amine"

XX WO9512808-A1.

XX PD 11-MAY-1995.

XX PF 26-OCT-1994; 94MO-US012270.

XX PR 01-NOV-1993; 93US-00146504.

XX PA (NANO-) NANOGEN INC.

XX PI Heller MJ, Tu E;

XX DR WPI; 1995-185870/24.

XX New self-addressable electronic devices - used for multi-step and  
PT multiplex reactions such as DNA hybridisation(s), clinical diagnostics  
PT and bio/polymer synthesis.

XX Example 1; Page 41; 86pp; English.

XX The sequences represented by, AAQ90402-15 are synthetic DNA probes  
CC containing 5' amino termini. The sequences shown in AAQ90390-401 are  
CC synthetic DNA probes with 3' ribonucleoside termini. These sequences were  
CC specific for the polymorphisms of HLA gene dQa. The sequences were used  
CC in the device of the invention. This is a self-addressable electronic  
CC device (SAED) that can be used to carry out multi-step and multiplex  
CC reactions, such as nucleic acid hybridisations. The advantages of this  
CC method are that these reactions can be carried out with complete and  
CC precise electronic control, and that the rate, specificity and  
CC sensitivity of these reactions are greatly improved at micro-locations

XX Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 8

AA73704/C  
ID AA73704 standard; DNA; 20 BP.

XX AA73704;

XX 27-FEB-1998 (first entry)

DE PCR primer used to prepare probes for diagnosing Alzheimer's disease.

XX PCR primer BK33; Alzheimer's disease; probe; diagnosis; fluorochrome;  
KM yeast artificial chromosome library; YAC; chromosome 14; presentie; ss.

XX Synthetic.

XX FR2742758-A1.

XX PD 27-JUN-1997.

XX PF 28-OCT-1994; 94FR-00012941.

XX PR 28-OCT-1994; 94FR-00012941.

XX PA (ASFR-) ASSOC FR CONTRA MYOPATHIES ASSOC LOT.

XX Weissenbach J, Hellig R;

XX WPI; 1997-353201/33.

XX Probes for diagnosing Alzheimer's disease - hybridising with chromosome  
PT 14 segments cloned in yeast artificial chromosome library.

XX Example 1; Page 8; 21pp; French.

XX PCR primers AAT73703-4 were used to prepare probes (containing Alu  
CC repeats) for detecting a mutation in the locus of chromosome 14  
CC associated with a presenile form of Alzheimer's disease. Each of the  
CC probes hybridises with one of the two human chromosomal DNA segments  
CC cloned in the CEPH yeast artificial chromosome (YAC) library under the  
CC accession numbers YAC 93443 identifiable by genetic marker D14S251 and  
CC YAC 85445 (identifiable by genetic marker D14S76). The probes are useful  
CC for diagnosis of the form of Alzheimer's disease associated with  
CC chromosomes 14 by a method comprising making a preparation of metaphase  
CC slides; contacting the preparation under DNA hybridisation conditions with  
CC the pair of probes or with one of the probes and another probe that  
CC hybridises with YAC 905C2 from the same library, and detecting the  
CC hybridised probes and their relative positions on a significant number of  
CC pairs of chromosomes

XX Sequence 20 BP; 4 A; 4 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1844 GCTCAGCCCTGTAATCCAG 1863  
DB 20 GCTCAGCCCTGTAATCCAG 1

RESULT 9

AAT63648/C  
ID AAT63649 standard; DNA; 20 BP.

XX AAT63649;

XX 06-JUN-1997 (first entry)

DE Anti-HTLV antisense reference oligonucleotide HT.

KW antisense; complementary; tax gene; inhibit; HTLV-1;  
 KW human T-cell lymphotropic virus type 1; viral antigen expression; ss.  
 XX Synthetic.  
 OS JP09052898-A.  
 XX  
 XX  
 XX 25-FEB-1997.  
 PD  
 XX 09-AUG-1995; 95UP-00224606.  
 PF  
 XX 09-AUG-1995; 95UP-00224606.  
 PR  
 XX (SOYA-) SOYAKU GIUTSU KENKUSHO KK.  
 PA  
 XX WPI; 1997-197252/18.  
 DR  
 XX Anti-HTLV-1 anti-sense oligo:nucleotide - is complementary to region of  
 PT tax gene from human T-cell lymphotropic virus type 1 and inhibits viral  
 PT antigen expression.  
 PT  
 XX Example 1; Page 8; 10pp; Japanese.  
 PS  
 XX Oligonucleotides having a partial sequence consisting of at least 15  
 CC bases of AAT63641 (an antisense oligo complementary to a region of the  
 CC tax gene which can inhibit human T-cell lymphotropic virus type 1 (HTLV-  
 CC 1) viral antigen expression) are claimed. In an example, six antisense  
 CC oligos were designed, T1-T6 (AAT63650-55) and were compared to six oligos  
 CC derived from other regions of HTLV-1, i.e. S1 (splice junction), P1  
 CC (p21), R1 (rex), R1 (rex response element), B1 (env) and G1 (gag), four  
 CC reference oligonucleotides T1S (tax-sense), HC (dc20), HT (dt20)  
 CC (AAT63647-49) and a random 20mer (RAM) in a HTLV-1 virus antigen  
 CC expression inhibiting test. Oligonucleotide T1 gave the best results  
 SQ Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
 |||||  
 Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 10  
 AAV85582/C  
 ID AAV85582 standard; DNA; 20 BP.  
 XX  
 AC AAV85582;  
 XX  
 DT 10-FEB-1999 (first entry)  
 XX  
 DE LRP5 PCR primer Gp1 1P.  
 XX

KW LRP5; LDL-receptor related protein; LRP-3; IDDM; diagnosis; endocytosis;  
 KW insulin dependent diabetes mellitus; autoimmune disease;  
 KW glomerulonephritis; inflammation; viral infection; osteoporosis;  
 KW hypercholesterolemia; Alzheimer's disease; low density lipoprotein;  
 KW PCR primer; ss.  
 XX  
 XX Synthetic.  
 OS Homo sapiens.  
 XX  
 XX WO9846743-A1.  
 PV  
 XX 22-OCT-1998.  
 PD  
 XX 15-APR-1998; 98WO-GB001102.  
 PF  
 XX 15-APR-1997; 97US-0043553P.  
 PR  
 XX 05-JUN-1997; 97US-0048740P.  
 XX

PA (WELT) WELLCOME TRUST LTD.  
 PA (MERT) MERCK & CO INC.  
 XX  
 XX Todd JA, Hess JW, Caskey CT, Cox RD, Gerhold D, Hammond H;  
 PI Hey P, Kawaguchi Y, Merriman TR, Metzker ML, Nakagawa Y;  
 PI Phillips MS, Twells RCJ;  
 XX  
 DR WPI; 1998-594573/50.  
 XX  
 XX New isolated LDL-receptor related protein - used to develop products for  
 PT treating, e.g. elevated triglyceride levels, diabetes, autoimmune  
 PT disorder, inflammation or Alzheimer's disease.  
 PS  
 XX Claim 12; Page 98; 200pp; English.  
 XX

CC The present invention describes LRP5 (low density lipoprotein (LDL)  
 CC receptor related protein, previously designated LRP-3). AAV85552 to  
 CC AAV85586 represent PCR primer used for obtaining LRP5 cDNA. Nucleic acid  
 CC molecules (NAs) encoding LRP5 can be used for determining if an  
 CC individual is susceptible to insulin dependent diabetes mellitus (IDDM).  
 CC The NAs or proteins can be used for reducing triglyceride levels in the  
 CC serum of an individual. Therapies that affect LRP5 may also be useful in  
 CC the treatment of autoimmune diseases such as glomerulonephritis, diseases  
 CC and disorders involving disruption of endocytosis and/or antigen  
 CC presentation, cytokine clearance and/or inflammation, viral infection,  
 CC pathogenic bacterial toxin contamination, elevation of free fatty acids  
 CC or hypercholesterolemia, type 2 diabetes, osteoporosis, Alzheimer's  
 CC disease and cardiovascular disease. Products from the present invention  
 CC can also be used for detection, diagnosis and drug screening  
 XX  
 SQ Sequence 20 BP; 4 A; 6 C; 5 G; 5 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1901 CAGGAGTTTGAGACCGCCT 1920  
 |||||  
 Db 20 CAGGAGTTTGAGACCGCCT 1

RESULT 11  
 AAV34591  
 ID AAV34591 standard; DNA; 20 BP.  
 XX  
 AC AAV34591;  
 XX  
 DT 25-AUG-1998 (first entry)  
 XX  
 DE M. vaccae antigenic sequence hybridising oligo AD12.  
 XX

KW Mycobacterium vaccae; antigen; therapy; prevention; cytokine production;  
 KW M. avium; M. tuberculosis; immune response enhancer; cell proliferation;  
 KW mycobacteria infection; vaccine; cancer; ss.  
 XX  
 OS Synthetic.  
 OS Mycobacterium vaccae.  
 XX  
 XX WO9808542-A2.  
 PN  
 XX 05-MAR-1998.  
 PD  
 XX 28-AUG-1997; 97WO-NZ000105.  
 PF  
 XX 29-AUG-1996; 96US-00705347.  
 PR  
 XX 12-JUN-1997; 97US-00873970.  
 XX

PA (GENE-) GENESIS RES & DEV CORP.  
 XX  
 XX Tan P, Hiyama J, Visser E, Skinner MA, Scott LM, Prestidge RL;  
 PI WPI; 1998-216926/19.  
 DR  
 XX

PT Mycobacterium vaccae polypeptides - used to develop products for use in  
 PT detection, therapy and prevention of mycobacteria infections or as immune  
 PT response enhancers.  
 XX  
 PS Example 8; Page 99; 153pp; English.  
 CC This oligonucleotide is used in the DNA cloning strategies of the  
 CC Mycobacterium vaccae antigens. The invention provides M. vaccae  
 CC polypeptides that comprise an immunogenic portion of a soluble M. vaccae  
 CC antigen, or a variant, where the antigen induces an immune response in  
 CC patients previously exposed to a mycobacterium. Such M. vaccae  
 CC polypeptides can be used in methods for enhancing non-specific immune  
 CC response. The methods and products can be used for the detection,  
 CC treatment and prevention of infectious diseases caused by mycobacteria  
 CC such as M. vaccae, M. avium or M. tuberculosis. The products also have  
 CC the ability to induce cell proliferation and cytokine production (e.g.  
 CC interferon-gamma and interleukin-12 production) in T cells, NK cells, B  
 CC cells, or macrophages. They can be used for enhancing immune responses  
 CC for use in vaccines or immunotherapy of infectious diseases and cancers  
 XX  
 SQ Sequence 20 BP; 0 A; 0 C; 0 G; 0 T; 0 U; 0 Other;  
 Query Match 0.7%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1799 AAAAAAAAAAAAAAAAAA 1818  
 Db 1 AAAAAAAAAAAAAAAAAA 20  
 RESULT 12  
 AAT86606/c  
 ID AAT86606 standard; DNA; 20 BP.  
 XX  
 AC AAT86606;  
 XX  
 DT 04-JUN-1998 (first entry)  
 XX  
 DE Oligonucleotide separated by capillary affinity gel electrophoresis.  
 XX  
 KW Capillary affinity gel electrophoresis; separation; polymer-gel;  
 KM polyacrylamide; ss.  
 XX  
 OS Synthetic.  
 OS  
 PN WO9745721-A1.  
 PD  
 PD 04-DEC-1997.  
 XX  
 PF 23-MAY-1997; 97WO-EP002647.  
 XX  
 PR 24-MAY-1996; 96CH-00001320.  
 XX  
 PA (NOVS ) NOVARTIS AG.  
 XX  
 XX Muscate A, Paulus A, Natt F;  
 PI  
 PI WPI; 1998-041763/04.  
 DR  
 XX  
 PT Separation of electrically charged target molecules - by capillary  
 PT affinity gel electrophoresis using polymer-gel to which receptors for  
 PT target molecules are bound.  
 XX  
 PS Example D3; Page 25; 41pp; English.  
 XX  
 CC A mixture of oligonucleotides (AAT86604-7) were separated by a new  
 CC process using capillary affinity gel electrophoresis. The invention  
 CC relates to selective separation of electrically charged target molecules  
 CC in an analytical mixture. It comprises capillary affinity gel  
 CC electrophoresis using a capillary tube which is at least partly filled  
 CC with a polymer gel. Receptors for target molecules are covalently bound  
 CC to the polymer. An electric field of at least 50 volts/cm is applied. The

CC capillary tube is charged with the analytical mixture. In a first  
 CC separation stage, the target molecules in the mixture are bound to the  
 CC receptors and the remaining components are eluted, optionally whilst  
 CC splitting open. In a second stage, the elution conditions are changed,  
 CC optionally in stages, so that the affinity of the target molecules for  
 CC the receptor is eliminated and the target molecules are eluted and  
 CC detected, optionally whilst splitting open. The process is useful for  
 CC selective separation and/or determination of charged organic compounds,  
 CC such as oligonucleotides, peptides or carbohydrates. It may be used, e.g.  
 CC for isolation of specific proteins and DNA molecules, purification of  
 CC antibodies, analysis of antisense compounds or screening for enzyme  
 CC inhibitors. The process achieves higher resolution and selectivity than  
 CC prior art processes, especially in the case of complex biological  
 CC samples. The derivatised polymers may be synthesised specifically using  
 CC standard methods  
 XX  
 SQ Sequence 20 BP; 0 A; 0 C; 0 G; 20 T; 0 U; 0 Other;  
 Query Match 0.7%; Score 20; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1799 AAAAAAAAAAAAAAAAAA 1818  
 Db 20 AAAAAAAAAAAAAAAAAA 1  
 RESULT 13  
 AA237736  
 ID AA237736 standard; DNA; 20 BP.  
 XX  
 AC AA237736;  
 XX  
 DT 07-JAN-2000 (first entry)  
 XX  
 DE Human mdm2 phosphorothioate oligodeoxynucleotide #266.  
 XX  
 KW Human mdm2 gene; proliferation; tumour; phosphorothioate; p53; cancer;  
 KW antisense; modulation; oligonucleotide; expression; inhibition;  
 KW hyperproliferation; blood cancer; brain cancer; breast cancer;  
 KW lung cancer; soft tissue cancer; psoriasis; fibrosis; atherosclerosis;  
 KW restenosis; ss.  
 XX  
 OS Synthetic.  
 OS  
 OS Homo sapiens.  
 XX  
 XX  
 XX WO9949065-A1.  
 XX  
 PD 30-SEP-1999.  
 XX  
 PF 26-MAR-1999; 99WO-US006702.  
 XX  
 PR 26-MAR-1998; 98US-00048810.  
 XX  
 XX (ISIS-) ISIS PHARM INC.  
 XX  
 XX Miraglia LJ, Nero P, Graham MJ, Cowsett LM;  
 PI  
 PI WPI; 1999-610754/52.  
 DR  
 XX  
 PT New antisense compounds used to treat eg. hyperproliferative conditions.  
 PT  
 XX  
 PS Example 9; Page 55; 157pp; English.  
 XX  
 CC AA237473-237738 represent human mdm2 phosphorothioate oligonucleotides.  
 CC AA237471, AA237472, AA237739, AA237740 and AA237741 are used in the  
 CC exemplification of the present invention. The present invention describes  
 CC novel nucleotide antisense compounds, targeted to the 5' untranslated  
 CC translation termination codon, or 3' untranslated region of a nucleic  
 CC acid encoding human mdm2, that modulates expression of human mdm2. The  
 CC oligonucleotides mediate their effect by antisense inhibition of  
 CC hyperproliferative gene expression. The antisense compound is used to

CC treat an animal having a disease or condition associated with mdm2.  
CC particularly a hyperproliferative condition, more particularly cancer,  
CC especially of the blood, brain, breast, lung or soft tissue, or  
CC psoriasis, fibrosis, atherosclerosis or restenosis

XX Sequence 20 BP; 3 A; 7 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1858 TCCGACGACTTTGGAGGCC 1877

Db 1 TCCGACGACTTTGGAGGCC 20

RESULT 14

AAZ37737  
ID AAZ37737 standard; DNA; 20 BP.

XX AAZ37737;

DT 07-JAN-2000 (first entry)

DE Human mdm2 phosphorothioate oligodeoxynucleotide #267.

XX Human mdm2 gene; proliferation; tumour; phosphorothioate; p53; cancer;  
XX antisense; modulation; oligonucleotide; expression; inhibition;  
XX hyperproliferation; blood cancer; brain cancer; breast cancer;  
XX lung cancer; soft tissue cancer; psoriasis; fibrosis; atherosclerosis;  
XX restenosis; ss.

XX Synthetic.

OS Homo sapiens.

XX MO9949065-A1.

XX 30-SEP-1999.

XX 26-MAR-1999; 99WO-US006702.

XX 26-MAR-1999; 98US-00048810.

XX (ISIS-) ISIS PHARM INC.

XX Miraglia LJ, Nero P, Graham MJ, Monia BP, Cowest LM;

XX WPI; 1999-610754/52.

XX New antisense compounds used to treat eg. hyperproliferative conditions.

XX Example 9; Page 55; 157pp; English.

CC AAZ37473-227728 represent human mdm2 phosphorothioate oligonucleotides.  
CC AAZ37471, AAZ37472, AAZ37739, AAZ37740 and AAZ37741 are used in the  
CC exemplification of the present invention. The present invention describes  
CC novel nucleotide antisense compounds, targeted to the 5' untranslated,  
CC translation termination codon, or 3' untranslated region of a nucleic  
CC acid encoding human mdm2, that modulates expression of human mdm2. The  
CC oligonucleotides mediate their effect by antisense inhibition of  
CC hyperproliferative gene expression. The antisense compound is used to  
CC treat an animal having a disease or condition associated with mdm2,  
CC particularly a hyperproliferative condition, more particularly cancer,  
CC especially of the blood, brain, breast, lung or soft tissue, or  
CC psoriasis, fibrosis, atherosclerosis or restenosis

XX Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1851 CCTGTAATCCGACGACTTTG 1870

Db 1 CCTGTAATCCGACGACTTTG 20

RESULT 15  
AAZ27533/c  
ID AAZ27533 standard; RNA; 20 BP.

XX AAZ27533;

DT 27-MAY-1999 (first entry)

DE Synthetic RNA sequence produced by the method of the invention.

XX silyloxymethyl; phosphonate; silyloxymethyl halide; diagnosis; ss;  
XX cyanoethyl phosphoramidate coupling; isomerisation; steric hindrance.

XX Synthetic.

XX WO9909044-A1.

XX 25-FEB-1999.

XX 17-AUG-1998; 98WO-EP005215.

XX 18-AUG-1997; 97CH-00001931.

XX (PITS/) PITSCH S.

XX (WEIS/) WEISS P. A.

XX (JENNY/) JENNY L.

XX Pitsch S, Weiss PA, Jenny L;

XX WPI; 1999-180963/15.

XX 2-Silyloxymethyl ribonucleosides and their phosphonate derivatives - have  
XX high purity, use in machine synthesis of ribonucleic acids, enable longer  
XX oligonucleotide chain construction, and larger amounts.

XX Example 6; Page 25; 38pp; English.

CC The invention relates to silyloxymethyl protected D- or L-ribonucleosides  
CC and their phosphonates (I), and silyloxymethyl halides (II). (I) are  
CC intermediates for synthesis of RNA-oligonucleotides with predetermined  
CC nucleotide sequence, particularly by machine synthesis. The groups  
CC specified above, apart from those on silyl, are those particularly for  
CC the cyanoethyl phosphoramidate coupling. Uses of the oligonucleotide  
CC products in diagnosis, therapy, and as research tools, are well known,  
CC and are not dealt with in detail. (II) is an intermediate for (I). The  
CC silyloxymethyl halide reagent is easy to prepare, and yields are high.  
CC Introduction of the silyloxymethyl group into the ribonucleoside is  
CC simple and rapid, and the acetal bond formed does not migrate,  
CC eliminating particularly the prior art problem of 2' to 3' isomerisation.  
CC The methylatedoxy group spacer between the silyl group and nucleoside  
CC ring results in less steric hindrance than bulky direct silyloxy  
CC linkages, enabling first, a range of choices for the silyl substituents,  
CC to provide, e.g., acid or base stability; and second, higher yields in  
CC coupling. Purer products are therefore obtained than in prior art,  
CC enabling larger quantities and longer chains of oligoribonucleotides to  
CC be synthesised successfully, and in shorter times

XX Sequence 20 BP; 0 A; 0 C; 0 G; 0 T; 20 U; 0 Other;

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818

Db 20 AAAAAAAAAAAAAAAAAA 1

Search completed: March 7, 2004, 15:07:19

Mon Mar 8 10:28:42 2004

us-09-966-880a-7.Oligo.rng

Page 8

Job time : 113 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 14:32:38 ; Search time 216 Seconds

(without alignments)  
7240.054 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818  
Sequence: 1 agagaaccatcatatga.....aaaaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 353258

Minimum DB seq length: 0  
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database :

Issued\_Patents.NA.\*  
1: /cgn2\_6/prodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	0.7	20	1 US-08-146-504-16	Sequence 16, Appl
C 2	20	0.7	20	2 US-08-379-593-5	Sequence 5, Appl
C 3	20	0.7	20	2 US-08-725-976-16	Sequence 16, Appl
C 4	20	0.7	20	2 US-08-997-080-83	Sequence 83, Appl
C 5	20	0.7	20	2 US-08-997-362-83	Sequence 83, Appl
C 6	20	0.7	20	2 US-08-965-780-1	Sequence 1, Appl
C 7	20	0.7	20	2 US-08-873-970-83	Sequence 83, Appl
C 8	20	0.7	20	3 US-08-765-340-96	Sequence 96, Appl
C 9	20	0.7	20	3 US-09-095-855-83	Sequence 83, Appl
C 10	20	0.7	20	3 US-09-407-675-1	Sequence 1, Appl
C 11	20	0.7	20	3 US-09-280-805-26	Sequence 26, App
C 12	20	0.7	20	3 US-09-280-805-26	Sequence 26, App
C 13	20	0.7	20	3 US-09-250-075-1	Sequence 1, Appl
C 14	20	0.7	20	3 US-09-173-936B-14	Sequence 14, Appl
C 15	20	0.7	20	3 US-09-454-704A-13	Sequence 13, Appl
C 16	20	0.7	20	4 US-09-488-856A-62	Sequence 62, Appl
C 17	20	0.7	20	4 US-09-324-542-83	Sequence 83, Appl
C 18	20	0.7	20	4 US-09-203-426-83	Sequence 83, Appl
C 19	20	0.7	20	4 US-09-619-103-26	Sequence 26, Appl
C 20	20	0.7	20	4 US-09-726-096A-1	Sequence 1, Appl
C 21	20	0.7	20	4 US-09-733-294A-82	Sequence 82, Appl
C 22	20	0.7	20	4 US-09-603-830-55	Sequence 55, Appl
C 23	20	0.7	20	4 US-09-976-978A-55	Sequence 55, Appl
C 24	20	0.7	20	4 US-09-060-299-78	Sequence 78, Appl
C 25	20	0.7	20	4 US-09-402-923A-78	Sequence 78, Appl
C 26	20	0.7	20	4 US-09-679-299A-70	Sequence 70, Appl
C 27	20	0.7	20	4 US-09-344-260A-10	Sequence 10, Appl

28	20	0.7	20	4 US-09-961-949A-55	Sequence 55, Appl
29	20	0.7	20	4 US-09-966-491A-55	Sequence 55, Appl
30	20	0.7	20	4 US-09-957-113A-55	Sequence 55, Appl
31	20	0.7	20	4 US-09-966-112-55	Sequence 55, Appl
32	20	0.7	20	4 US-09-975-062A-55	Sequence 55, Appl
33	20	0.7	20	4 US-09-976-971A-55	Sequence 55, Appl
34	20	0.7	20	5 PCT-US93-07603-6	Sequence 6, Appl
C 35	19	0.7	19	1 US-08-629-939-10	Sequence 10, Appl
C 36	19	0.7	19	1 US-08-759-873-10	Sequence 10, Appl
C 37	19	0.7	19	1 US-08-756-728A-1	Sequence 1, Appl
C 38	19	0.7	19	2 US-08-469-852A-2	Sequence 2, Appl
C 39	19	0.7	19	3 US-08-271-882B-16	Sequence 16, Appl
C 40	19	0.7	19	3 US-08-295-509B-2	Sequence 2, Appl
C 41	19	0.7	19	3 US-09-234-237-1	Sequence 1, Appl
C 42	19	0.7	19	3 US-09-016-520-20	Sequence 20, Appl
C 43	19	0.7	19	3 US-09-016-520-21	Sequence 21, Appl
C 44	19	0.7	19	3 US-09-016-520-22	Sequence 22, Appl
C 45	19	0.7	19	3 US-09-016-520-23	Sequence 23, Appl

#### ALIGNMENTS

RESULT 1  
US-08-146-504-16/c  
Sequence 16, Application US/08146504  
Patent No. 5605662  
GENERAL INFORMATION:  
APPLICANT: Heller, Michael J.; and Tu, Eugene  
TITLE OF INVENTION: SELF-ADDRESSABLE SELF-ASSEMBLING  
TITLE OF INVENTION: MICROELECTRONIC SYSTEMS AND DEVICES FOR  
TITLE OF INVENTION: MOLECULAR BIOLOGICAL ANALYSIS AND  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 611 West Sixth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90017  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/146,504  
FILING DATE: No. 5605662ember 1, 1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 203/218  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 673510  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-146-504-16  
Query Match 0.7%; Score 20; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;



Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818

DB 20 AAAAAAAAAAAAAAAAAA 1

## RESULT 2

US-08-379-593-5/c

Sequence 5, Application US/08379593

Patent No. 5849480

GENERAL INFORMATION:

APPLICANT: Cros, Philippe

APPLICANT: Kufirst, Robin

APPLICANT: Battail, Nicole

APPLICANT: Piga, Nadia

TITLE OF INVENTION: HAPTEN ASSAY DEVICE AND USE THEREOF

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESSEE: OLIFF & BERRIDGE

STREET: 700 South Washington Street, Suite 300

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Floppy disk, 1.44M storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/379,593

FILING DATE: 02-FEB-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Berridge, William P.

REGISTRATION NUMBER: 30,024

REFERENCE/DOCKET NUMBER: WPB 36056

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6400

TELEFAX: 703-836-2787

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "SYNTHETIC DNA"

FEATURE:

OTHER INFORMATION: consists of nucleosides with an alpha anomer and carries

US-08-379-593-5

QY 1799 AAAAAAAAAAAAAAAAAA 1818

DB 20 AAAAAAAAAAAAAAAAAA 1

## RESULT 3

US-08-725-976-16/c

Sequence 16, Application US/08725976

Patent No. 5929208

GENERAL INFORMATION:

APPLICANT: Heller, Michael J. and Ty, Eugene

APPLICANT: Battail, Nicole

TITLE OF INVENTION: METHODS FOR ELECTRONIC SYNTHESIS OF POLYMERS

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 MB

COMPUTER: IBM compatible

OPERATING SYSTEM: WINDOWS (VERSION 3.0)

SOFTWARE: WordPerfect (Version 6.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/725,976

FILING DATE: October 4, 1996

CLASSIFICATION: 422

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application

PRIOR APPLICATION DATA: described below:

APPLICATION NUMBER: 08/146,504

FILING DATE: No. 5929208ember 1, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Murphy, David B.

REGISTRATION NUMBER: 31,125

REFERENCE/DOCKET NUMBER: 222/211

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 20

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-725-976-16

QY 1799 AAAAAAAAAAAAAAAAAA 1818

DB 20 AAAAAAAAAAAAAAAAAA 1

## RESULT 4

US-08-997-080-83

Sequence 83, Application US/08997080

Patent No. 5968524

GENERAL INFORMATION:

APPLICANT: WATSON, JAMES D.

APPLICANT: TAN, PAUL L.J.

TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-

NUMBER OF SEQUENCES: 194

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/997,080

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet

REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0563  
TELEX:  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Other  
US-08-997-080-83

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 5  
US-08-997-362-83  
Sequence 83, Application US/08997362  
Patent No. 5985287  
GENERAL INFORMATION:  
APPLICANT: Tan, Paul  
APPLICANT: Hiya, Jun  
APPLICANT: Visser, Elizabeth  
APPLICANT: Skinner, Margot  
APPLICANT: Scott, Linda  
APPLICANT: Prestidge, Rose  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR  
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS  
TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS  
NUMBER OF SEQUENCES: 194  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Ann W. Speckman  
STREET: 2601 Elliott Avenue, Suite 4185  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997.362  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/673.970  
FILING DATE: June 12, 1997  
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705.347  
FILING DATE: August 29, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleach, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1002c2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0563  
TELEX:  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: Other  
US-08-997-362-83

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 6  
US-08-965-780-1/c  
Sequence 1, Application US/08965780  
Patent No. 5986084  
GENERAL INFORMATION:  
APPLICANT: Pitsch, Stefan  
APPLICANT: Weiss, Patrick A.  
APPLICANT: Jenny, Luzi  
TITLE OF INVENTION: RIBONUCLEOSIDE-DERIVATIVE AND METHOD FOR  
PREPARING THE SAME  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KUBOVCIK & KUBOVCIK  
STREET: 900 17th Street, N.W., Suite 990  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/965.780  
FILING DATE: 07-NOV-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 01931/97  
FILING DATE: 18-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubovcik, Ronald J.  
REGISTRATION NUMBER: 25,401  
REFERENCE/DOCKET NUMBER: FREI-002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-887-9023  
TELEFAX: 202-887-9093  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Other nucleic acid  
DESCRIPTION: /desc = "oligoribonucleotide"  
US-08-965-780-1

Query Match 0.7%; Score 20; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 7  
US-08-873-970-83  
Sequence 83, Application US/08873970  
Patent No. 6001361  
GENERAL INFORMATION:

APPLICANT: Tan, Paul  
APPLICANT: Hiwama, Jun  
APPLICANT: Visser, Elizabeth  
APPLICANT: Skinner, Margot  
APPLICANT: Scott, Linda  
APPLICANT: Prestidge, Ross  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR  
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Ann W. Speckman  
STREET: 2601 Elliott Avenue, Suite 4185  
City: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/873,970  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/705,347  
FILING DATE: 29-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1002C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0563  
TELEX:  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other  
US-08-873-970-83

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
1 AAAAAAAAAAAAAAAAAA 20  
DB

RESULT 8  
US-08-765-340-96/C  
Sequence 96, Application US/08765340  
Patent No. 6150092  
GENERAL INFORMATION:  
APPLICANT: UCHIDA, K.,  
APPLICANT: UCHIDA, T.,  
APPLICANT: TANAKA, Y.,  
APPLICANT: MATSUDA, Y.,  
APPLICANT: KONDO, S.,  
TITLE OF INVENTION: AN ANTISENSE NUCLEIC ACID  
TITLE OF INVENTION: COMPOUND  
NUMBER OF SEQUENCES: 185  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA

ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version  
SOFTWARE: #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/765,340  
FILING DATE: 23-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 145146/94  
FILING DATE: 27-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 311130/94  
FILING DATE: 21-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: SERUNIAN, LESLIE  
REGISTRATION NUMBER: 35,353  
REFERENCE/DOCKET NUMBER: 1452-4005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "synthetic DNA"  
US-08-765-340-96

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
20 AAAAAAAAAAAAAAAAAA 1  
DB

RESULT 9  
US-09-095-855-83  
Sequence 83, Application US/09095855  
Patent No. 6160093  
GENERAL INFORMATION:  
APPLICANT: Tan, Paul  
APPLICANT: Visser, Elizabeth  
APPLICANT: Skinner, Margot  
APPLICANT: Prestidge, Ross  
TITLE OF INVENTION: Compounds and Methods for  
Treatment and Diagnosis of Mycobacterial Infections  
NUMBER OF SEQUENCES: 208  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Ann W. Speckman  
STREET: 2601 Elliott Avenue, Suite 4185  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/095,855  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/705,347  
FILING DATE: 29-AUG-1996

APPLICATION NUMBER: 08/873,970  
FILING DATE: 12-JUN-1997  
APPLICATION NUMBER: 08/997,362  
FILING DATE: 23-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1002c3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0563  
TELEX:  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Other  
US-09-95-855-83

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
1 AAAAAAAAAAAAAAAAAA 20

Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 10  
US-09-407-675-1  
Sequence 1, Application US/09407675  
Patent No. 6169176  
GENERAL INFORMATION:  
APPLICANT: Bruice, Thomas C.  
APPLICANT: Arya, Dev P.  
TITLE OF INVENTION: DEOXYNUCLEIC ALKYL THIUREA COMPOUNDS AND USES THEREOF  
FILE REFERENCE: 30448.65U802  
CURRENT APPLICATION NUMBER: US/09/407,675  
CURRENT FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: 09/347,443  
PRIOR FILING DATE: 1999-07-02  
PRIOR APPLICATION NUMBER: 60/091,481  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/111,800  
PRIOR FILING DATE: 1998-12-11  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: Oligo 1  
US-09-407-675-1

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
1 AAAAAAAAAAAAAAAAAA 20

Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 11  
US-09-280-805-266  
Sequence 266, Application US/09280805  
Patent No. 6184212  
GENERAL INFORMATION:  
APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.  
APPLICANT: Graham, Brett P. Monia

TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDM2  
NUMBER OF SEQUENCES: 271  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Jane Massey Licata  
STREET: 66 East Main Street  
CITY: Marlton  
STATE: NJ  
COUNTRY: U.S.A.  
ZIP: 08053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
COMPUTER: IBM PC  
OPERATING SYSTEM: WINDOWS 95  
SOFTWARE: WORDPERFECT 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/280,805  
FILING DATE: herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/048,810  
FILING DATE: March 26, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Licata, Jane Massey  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: ISPH-0346  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-810-1515  
TELEFAX: 609-810-1454  
INFORMATION FOR SEQ ID NO: 266:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
ANTI-SENSE: yes  
US-09-280-805-266

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TCCGACGACTTGGAGGCC 1877  
1 TCCGACGACTTGGAGGCC 20

Db 1 TCCGACGACTTGGAGGCC 20

RESULT 12  
US-09-280-805-267  
Sequence 267, Application US/09280805  
Patent No. 6184212  
GENERAL INFORMATION:  
APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.  
APPLICANT: Graham, Brett P. Monia  
TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDM2  
NUMBER OF SEQUENCES: 271  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Jane Massey Licata  
STREET: 66 East Main Street  
CITY: Marlton  
STATE: NJ  
COUNTRY: U.S.A.  
ZIP: 08053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
COMPUTER: IBM PC  
OPERATING SYSTEM: WINDOWS 95  
SOFTWARE: WORDPERFECT 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/280,805  
FILING DATE: herewith  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/048,810  
FILING DATE: March 26, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Licata, Jane Massey  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: ISPH-0346  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-810-1515  
TELEFAX: 609-810-1454  
INFORMATION FOR SEQ ID NO: 267:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
ANTI-SENSE: Yes  
US-09-280-805-267

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1851 CCTGTATCCCGACACTTGG 1870  
DB 1 CCTGTATCCCGACACTTGG 20

RESULT 13  
US-09-250-075-1/c  
Sequence 1, Application US/09250075  
Patent No. 6207819  
GENERAL INFORMATION:  
APPLICANT: Manoharan, Muthiah  
APPLICANT: Maier, Martin A  
TITLE OF INVENTION: Compounds Processes And Intermediates For Synthesis Of  
TITLE OF INVENTION: Mixed Backbone Oligomeric Compounds  
FILE REFERENCE: ISIS3299  
CURRENT APPLICATION NUMBER: US/09/250,075  
CURRENT FILING DATE: 1999-02-12  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(19)  
OTHER INFORMATION: 2'-methoxyethoxy (MOE)  
OTHER INFORMATION: Description of Artificial Sequence: No. 6207819e1  
OTHER INFORMATION: Sequence  
US-09-250-075-1

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 14  
US-09-173-936B-14/c  
Sequence 14, Application US/09173936B  
Patent No. 6238865  
GENERAL INFORMATION:  
APPLICANT: Zhen, Huang; Szostrak, Jack W.  
TITLE OF INVENTION: A Simple and Efficient Method to Label and Modify 3'-  
Terminal  
of RNA Using DNA Polymerase and a Synthetic Template with D  
Nucleotides

NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cohen, Pontani, Lieberman & Pavana  
STREET: 551 Fifth Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10176  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.50 Inch Diskette  
COMPUTER: IBM-MS  
OPERATING SYSTEM: Window 95  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/173,936B  
FILING DATE: 16-Oct-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/063,757  
FILING DATE: 17-Oct-1997  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-173-936B-14

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 15  
US-09-454-704A-13  
Sequence 13, Application US/09454704A  
Patent No. 6274321  
GENERAL INFORMATION:  
APPLICANT: Blumberg, Bruce  
TITLE OF INVENTION: High Throughput Functional Screening of  
TITLE OF INVENTION: CDNs  
FILE REFERENCE: P-UC 3662  
CURRENT APPLICATION NUMBER: US/09/454,704A  
CURRENT FILING DATE: 1999-12-03  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: cDNA  
US-09-454-704A-13

Query Match 0.7%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 1 AAAAAAAAAAAAAAAAAA 20

Search completed: March 7, 2004, 20:11:12  
Job time : 218 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 18:09:14 ; Search time 963 Seconds

(without alignments)  
10702.641 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818  
Sequence: 1 gagagaccctcatcattga.....aaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2421054 seqs, 1828716029 residues

Word size : 0

Total number of hits satisfying chosen parameters: 452776

Minimum DB seq length: 0  
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	0.7	20	US-09-752-983-266	Sequence 266, App
2	20	0.7	20	US-09-752-983-267	Sequence 267, App
3	20	0.7	20	US-09-733-294A-82	Sequence 82, App
4	20	0.7	20	US-09-973-788A-55	Sequence 55, App
5	20	0.7	20	US-09-973-638A-55	Sequence 55, App
6	20	0.7	20	US-09-974-007-55	Sequence 55, App
7	20	0.7	20	US-09-976-617A-55	Sequence 55, App
8	20	0.7	20	US-09-961-949A-55	Sequence 55, App
9	20	0.7	20	US-09-760-500A-55	Sequence 55, App
10	20	0.7	20	US-09-967-409A-55	Sequence 55, App
11	20	0.7	20	US-09-975-062A-55	Sequence 55, App
12	20	0.7	20	US-09-976-378A-55	Sequence 55, App
13	20	0.7	20	US-09-976-577-55	Sequence 55, App
14	20	0.7	20	US-09-771-554-5	Sequence 5, Appl
15	20	0.7	20	US-09-966-312-55	Sequence 55, Appl

16	20	0.7	20	US-09-927-777A-55	Sequence 55, Appl
17	20	0.7	20	US-09-927-777A-70	Sequence 70, Appl
18	20	0.7	20	US-09-966-491A-55	Sequence 55, Appl
19	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
20	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
21	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
22	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
23	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
24	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
25	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
26	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
27	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
28	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
29	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
30	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
31	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
32	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
33	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
34	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
35	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
36	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
37	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
38	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
39	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
40	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
41	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
42	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
43	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
44	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl
45	20	0.7	20	US-09-976-971A-55	Sequence 55, Appl

## ALIGNMENTS

RESULT 1  
US-09-752-983-266  
Sequence 266, Application US/09752983  
Patent No. US2001016575A1  
GENERAL INFORMATION:  
APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.  
APPLICANT: Graham, Brett P. Monia  
TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDN2  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 271  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Jane Massey Licata  
STREET: 66 East Main Street  
CITY: Marlton  
STATE: NJ  
COUNTRY: U.S.A.  
ZIP: 08053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
COMPUTER: IBM PC  
OPERATING SYSTEM: WINDOWS 95  
SOFTWARE: WORDPERFECT 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/752,983  
CLASSIFICATION:  
FILING DATE: 02-Jan-2001  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 09/280,805  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Licata, Jane Massey  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: ISPH-0346  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-810-1515  
TELEFAX: 609-810-1454  
INFORMATION FOR SEQ ID NO: 266:  
SEQUENCE CHARACTERISTICS:

LENGTH: 20 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
ANTI-SENSE: Yes  
US-09-752-983-266

Query Match  
Best Local Similarity 0.7%; Score 20; DB 9; Length 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TCCGACACTTGGGAGGCC 1877  
DB 1 TCCGACACTTGGGAGGCC 20

## RESULT 2

US-09-752-983-267  
Sequence 267, Application US/09752983  
Patent No. US20010016575A1

GENERAL INFORMATION:  
APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.  
APPLICANT: Graham, Brett P. Monia  
TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDW2  
TITLE OF INVENTION: EXPRESSION  
NUMBER OF SEQUENCES: 271  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Jane Massey Licata  
STREET: 66 East Main Street  
CITY: Marlton  
STATE: NJ  
COUNTRY: U.S.A.  
ZIP: 08053

COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE  
COMPUTER: IBM PC  
OPERATING SYSTEM: WINDOWS 95  
SOFTWARE: WORDPERFECT 6.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/752,983  
FILING DATE: 02-Jan-2001  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/280,805  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Licata, Jane Massey  
REGISTRATION NUMBER: 32,257  
REFERENCE/DOCKET NUMBER: 1SPH-0346  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-810-1515  
TELEFAX: 609-810-1454

INFORMATION FOR SEQ ID NO: 267:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
ANTI-SENSE: Yes  
US-09-752-983-267

Query Match  
Best Local Similarity 0.7%; Score 20; DB 9; Length 20;  
Matches 20; Conservative 100.0%; Pred. No. 55;  
Mismatches 0; Indels 0; Gaps 0;

QY 1851 CCTGTATCCCGACACTTG 1870  
DB 1 CCTGTATCCCGACACTTG 20

## RESULT 3

US-09-733-294A-82/C  
Sequence 82, Application US/09733294A

Patent No. US20020045588A1  
GENERAL INFORMATION:  
APPLICANT: Brett P. Monia  
APPLICANT: William M. Gaarde  
APPLICANT: Susan M. Freier  
APPLICANT: Edward V. Manciewicz  
TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION  
FILE REFERENCE: 1SPH-0527  
CURRENT APPLICATION NUMBER: US/09/733,294A  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: 09/572,423  
PRIOR FILING DATE: 2000-05-16  
NUMBER OF SEQ ID NOS: 108  
SEQ ID NO 82  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Antisense Oligonucleotide  
US-09-733-294A-82

Query Match  
Best Local Similarity 0.7%; Score 20; DB 9; Length 20;  
Matches 20; Conservative 100.0%; Pred. No. 55;  
Mismatches 0; Indels 0; Gaps 0;

QY 1846 TCACGCGCTGTATCCGACA 1865  
DB 20 TCACGCGCTGTATCCGACA 1

## RESULT 4

US-09-973-788A-55  
Sequence 55, Application US/09973788A  
Patent No. US20020127574A1

GENERAL INFORMATION:  
APPLICANT: Markin, Chad A.  
APPLICANT: Letsinger, Robert L.  
APPLICANT: Mucic, Robert C.  
APPLICANT: Storchoff, James J.  
APPLICANT: Elghanian, Robert  
APPLICANT: Talon, Thomas A.  
TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO  
TITLE OF INVENTION: AND USES THEREFOR  
FILE REFERENCE: 00-713-110  
CURRENT APPLICATION NUMBER: US/09/973,788A  
CURRENT FILING DATE: 2002-03-05  
PRIOR APPLICATION NUMBER: 09/603,830  
PRIOR FILING DATE: 2000-06-26  
PRIOR APPLICATION NUMBER: 09/344,667  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: 09/240,755  
PRIOR FILING DATE: 1999-01-29

PRIOR APPLICATION NUMBER: PCT/US97/12783  
PRIOR FILING DATE: 1997-07-21  
PRIOR APPLICATION NUMBER: 60/031,809  
PRIOR FILING DATE: 1996-07-29  
PRIOR APPLICATION NUMBER: 60/200,161  
PRIOR FILING DATE: 2000-04-26  
NUMBER OF SEQ ID NOS: 64  
SOFTWARE: Microsoft word 2000  
SEQ ID NO 55  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: random  
OTHER INFORMATION: synthetic sequence  
US-09-973-788A-55

Query Match  
Best Local Similarity 0.7%; Score 20; DB 9; Length 20;  
Matches 20; Conservative 100.0%; Pred. No. 55;  
Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 5  
US-09-973-638A-55  
; Sequence 55, Application US/09973638A  
; Patent No. US20020137070A1  
; GENERAL INFORMATION:  
; APPLICANT: Mirkin, Chad A.  
; APPLICANT: Letsinger, Robert L.  
; APPLICANT: Mucic, Robert C.  
; APPLICANT: Storchoff, James J.  
; APPLICANT: Elghanian, Robert  
; APPLICANT: Taton, Thomas A.  
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO  
; FILE REFERENCE: 00-713-19  
; CURRENT APPLICATION NUMBER: US/09/973,638A  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 09/603,830  
; PRIOR FILING DATE: 2000-06-26  
; PRIOR APPLICATION NUMBER: 09/344,667  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: 09/240,755  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: PCT/US97/12783  
; PRIOR FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/031,809  
; PRIOR FILING DATE: 1996-07-29  
; PRIOR APPLICATION NUMBER: 60/200,161  
; PRIOR FILING DATE: 2000-04-26  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: Microsoft Word 2000  
; SEQ ID NO 55  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: random  
; OTHER INFORMATION: synthetic sequence  
US-09-973-638A-55

Query Match 0.7%; Score 20; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 6  
US-09-974-007-55  
; Sequence 55, Application US/09974007  
; Patent No. US20020137071A1  
; GENERAL INFORMATION:  
; APPLICANT: Mirkin, Chad A.  
; APPLICANT: Letsinger, Robert L.  
; APPLICANT: Mucic, Robert C.  
; APPLICANT: Storchoff, James J.  
; APPLICANT: Elghanian, Robert  
; APPLICANT: Taton, Thomas A.  
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO  
; FILE REFERENCE: 00-713-18  
; CURRENT APPLICATION NUMBER: US/09/974,007  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 09/603,830  
; PRIOR FILING DATE: 2000-06-26  
; PRIOR APPLICATION NUMBER: 09/344,667  
; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: 09/240,755  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: PCT/US97/12783  
; PRIOR FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/031,809  
; PRIOR FILING DATE: 1996-07-29  
; PRIOR APPLICATION NUMBER: 60/200,161  
; PRIOR FILING DATE: 2000-04-26  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: Microsoft Word 2000  
; SEQ ID NO 55  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: random  
; OTHER INFORMATION: synthetic sequence  
US-09-974-007-55

Query Match 0.7%; Score 20; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 7  
US-09-976-617A-55  
; Sequence 55, Application US/09976617A  
; Patent No. US20020137072A1  
; GENERAL INFORMATION:  
; APPLICANT: Mirkin, Chad A.  
; APPLICANT: Letsinger, Robert L.  
; APPLICANT: Mucic, Robert C.  
; APPLICANT: Storchoff, James J.  
; APPLICANT: Elghanian, Robert  
; APPLICANT: Taton, Thomas A.  
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO  
; FILE REFERENCE: 00-713-124  
; CURRENT APPLICATION NUMBER: US/09/976,617A  
; CURRENT FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: 09/603,830  
; PRIOR FILING DATE: 2000-06-26  
; PRIOR APPLICATION NUMBER: 09/344,667  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: 09/240,755  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: PCT/US97/12783  
; PRIOR FILING DATE: 1997-07-21  
; PRIOR APPLICATION NUMBER: 60/031,809  
; PRIOR FILING DATE: 1996-07-29  
; PRIOR APPLICATION NUMBER: 60/200,161  
; PRIOR FILING DATE: 2000-04-26  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: Microsoft Word 2000  
; SEQ ID NO 55  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: random  
; OTHER INFORMATION: synthetic sequence  
US-09-976-617A-55

Query Match 0.7%; Score 20; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
Db 1 AAAAAAAAAAAAAAAAAA 20





```
RESULT 11
US-09-975-062A-55
; Sequence 55, Application US/09975062A
; Patent No. US20020155459A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Taton, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-111
; CURRENT APPLICATION NUMBER: US/09/975,062A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:random
US-09-975-062A-55

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 12
US-09-976-378A-55
; Sequence 55, Application US/09976378A
; Patent No. US20020155461A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Elghanian, Robert
; APPLICANT: Taton, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-125
; CURRENT APPLICATION NUMBER: US/09/976,378A
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
```

```

; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:random
US-09-976-378A-55

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 13
US-09-976-577-55
; Sequence 55, Application US/09976577
; Patent No. US20020155462A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Elghanian, Robert
; APPLICANT: Taton, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-120
; CURRENT APPLICATION NUMBER: US/09/976,577
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft Word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:random
US-09-976-577-55

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

RESULT 14
```

```

US-09-771-554-5/c
; Sequence 5, Application US/09771554
; Patent No. US20020155496A1
; GENERAL INFORMATION:
; APPLICANT: CHARLES, Marie Helene
; APPLICANT: FIGA, Nadia
; APPLICANT: BATTAIL-FOIROT, Nicole
; APPLICANT: VERON, Laurent
; APPLICANT: DELAIR, Thierry
; APPLICANT: MANDRAND, Bernard
; TITLE OF INVENTION: SATURATED AND UNSATURATED ABITANE DERIVATIVES, DERIVED CONJUGATE
; FILE REFERENCE: 108473
; CURRENT APPLICATION NUMBER: US/09/771,554
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: PCT/FR99/01846
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: FR 98/10084
; PRIOR FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-09-771-554-5

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 15
US-09-966-312-55
; Sequence 55, Application US/09966312
; Patent No. US20020164605A1
; GENERAL INFORMATION:
; APPLICANT: Mirkin, Chad A.
; APPLICANT: Letsinger, Robert L.
; APPLICANT: Mucic, Robert C.
; APPLICANT: Storchoff, James J.
; APPLICANT: Bighanian, Robert
; APPLICANT: Tator, Thomas A.
; TITLE OF INVENTION: NANOPARTICLES HAVING OLIGONUCLEOTIDES ATTACHED THERETO
; FILE REFERENCE: 00-713-15
; CURRENT APPLICATION NUMBER: US/09/966,312
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 09/603,830
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: 09/344,667
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/240,755
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: PCT/US97/12783
; PRIOR FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/031,809
; PRIOR FILING DATE: 1996-07-29
; PRIOR APPLICATION NUMBER: 60/200,161
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: Microsoft word 2000
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

```

; OTHER INFORMATION: Description of Artificial Sequence: random
; OTHER INFORMATION: synthetic sequence
US-09-966-312-55

Query Match          0.7%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
Db 1 AAAAAAAAAAAAAAAAAA 20

Search completed: March 7, 2004, 23:00:09
Job time : 964 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 14:48:48 ; Search time 9053 Seconds  
(without alignments)  
10952.666 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818

Sequence: 1 agagagaccatcatatctga.....aaaaaaaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 37577330 seqs, 17593059518 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2688236

Minimum DB seq length: 0

Maximum DB seq length: 20

Post-Processing: Listing first 45 summaries

Database :

Pending Patents NA Main:\*

1: /cgn2\_6/ptodata/2/pna/US1008\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US08\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US092B\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US093B\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US095\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US096\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US097\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US098D\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US099\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US099D\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US099E\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US099F\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US100A\_COMB.seq.\*

44: /cgn2\_6/ptodata/2/pna/US100B\_COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/US101A\_COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/US101B\_COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/US102A\_COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/US102B\_COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/US103A\_COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/US103B\_COMB.seq.\*  
51: /cgn2\_6/ptodata/2/pna/US104A\_COMB.seq.\*  
52: /cgn2\_6/ptodata/2/pna/US104B\_COMB.seq.\*  
53: /cgn2\_6/ptodata/2/pna/US105\_COMB.seq.\*  
54: /cgn2\_6/ptodata/2/pna/US107A\_COMB.seq.\*  
55: /cgn2\_6/ptodata/2/pna/US107B\_COMB.seq.\*  
56: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq.\*  
57: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq.\*  
58: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*  
59: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*  
60: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*  
61: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*  
62: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*  
63: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*  
64: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*  
65: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*  
66: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*  
67: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*  
68: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*  
69: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq.\*  
70: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*  
71: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*  
72: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*  
73: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*  
74: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*  
75: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*  
76: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*  
77: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*  
78: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*  
79: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*  
80: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq.\*  
81: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*  
82: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*  
83: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq.\*  
84: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq.\*  
85: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq.\*  
86: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq.\*  
87: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq.\*  
88: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq.\*  
89: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq.\*  
90: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq.\*  
91: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq.\*  
92: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq.\*  
93: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq.\*  
94: /cgn2\_6/ptodata/2/pna/US6037\_COMB.seq.\*  
95: /cgn2\_6/ptodata/2/pna/US6038\_COMB.seq.\*  
96: /cgn2\_6/ptodata/2/pna/US6039\_COMB.seq.\*  
97: /cgn2\_6/ptodata/2/pna/US6040\_COMB.seq.\*  
98: /cgn2\_6/ptodata/2/pna/US6041\_COMB.seq.\*  
99: /cgn2\_6/ptodata/2/pna/US6042\_COMB.seq.\*  
100: /cgn2\_6/ptodata/2/pna/US6043\_COMB.seq.\*  
101: /cgn2\_6/ptodata/2/pna/US6044\_COMB.seq.\*  
102: /cgn2\_6/ptodata/2/pna/US6045\_COMB.seq.\*  
103: /cgn2\_6/ptodata/2/pna/US6046\_COMB.seq.\*  
104: /cgn2\_6/ptodata/2/pna/US6047\_COMB.seq.\*  
105: /cgn2\_6/ptodata/2/pna/US6048\_COMB.seq.\*  
106: /cgn2\_6/ptodata/2/pna/US6049\_COMB.seq.\*  
107: /cgn2\_6/ptodata/2/pna/US6050\_COMB.seq.\*  
108: /cgn2\_6/ptodata/2/pna/US6051\_COMB.seq.\*  
109: /cgn2\_6/ptodata/2/pna/US6052\_COMB.seq.\*  
110: /cgn2\_6/ptodata/2/pna/US6053\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES



OTHER INFORMATION: Antisense Oligonucleotide  
PCT-US01-01411A-62

Query Match 0.7%; Score 20; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.9e+06;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1854 GTATCCCGACACTTTGGG 1873  
DB 20 GTATCCCGACACTTTGGG 1

RESULT 4  
PCT-US01-15774-82/c  
Sequence 82, Application PC/TUS0115774  
GENERAL INFORMATION:  
APPLICANT: Isis Pharmaceuticals, Inc.  
APPLICANT: Brett P. Monia  
APPLICANT: William Gaarde  
APPLICANT: Susan M. Freiler  
APPLICANT: Edward V. Wanciewicz  
TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION  
FILE REFERENCE: ISPH-0568  
CURRENT APPLICATION NUMBER: PCT/US01/15774  
CURRENT FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: 09/572,423  
PRIOR FILING DATE: 2000-05-16  
NUMBER OF SEQ ID NOS: 108  
SEQ ID NO 82  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Antisense Oligonucleotide  
PCT-US01-15774-82

Query Match 0.7%; Score 20; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.9e+06;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1846 TCACGCGTGTATCCGACA 1865  
DB 20 TCACGCGTGTATCCGACA 1

RESULT 5  
PCT-US01-30871-70/c  
Sequence 70, Application PC/TUS0130871  
GENERAL INFORMATION:  
APPLICANT: Isis Pharmaceuticals, Inc.  
APPLICANT: Vickie L. Brown-Driver  
APPLICANT: Hong Zhang  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 6 EXPRESSION  
FILE REFERENCE: RISP-0192  
CURRENT APPLICATION NUMBER: PCT/US01/30871  
CURRENT FILING DATE: 2001-10-03  
PRIOR APPLICATION NUMBER: 09/679,299  
PRIOR FILING DATE: 2000-10-04  
NUMBER OF SEQ ID NOS: 164  
SEQ ID NO 70  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Antisense Oligonucleotide  
PCT-US01-30871-70

Query Match 0.7%; Score 20; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.9e+06;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1852 CTGTATCCCGACACTTTGG 1871

DB 20 CTGTATCCCGACACTTTGG 1

RESULT 6  
PCT-US01-32233A-23/c  
Sequence 23, Application PC/TUS0132233A  
GENERAL INFORMATION:  
APPLICANT: Phyllos, Inc.  
TITLE OF INVENTION: PROTEIN SCAPFOLDS FOR ANTIBODY MIMICS  
FILE REFERENCE: 50036/021M04  
CURRENT APPLICATION NUMBER: PCT/US01/32233A  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: US 09/688,566  
PRIOR FILING DATE: 2000-10-16  
NUMBER OF SEQ ID NOS: 202  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 23  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Homo sapiens  
PCT-US01-32233A-23

Query Match 0.7%; Score 20; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.9e+06;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 7  
PCT-US01-44737-36/c  
Sequence 36, Application PC/TUS0144737  
GENERAL INFORMATION:  
APPLICANT: Weterings, Koen  
APPLICANT: Apuya, Nestor R.  
APPLICANT: Tatarianova, Tatiana  
APPLICANT: Goldberg, Robert B.  
APPLICANT: The Regents of the University of California  
APPLICANT: Ceres, Inc.  
TITLE OF INVENTION: Polynucleotides Useful for Modulating Transcription  
FILE REFERENCE: 023070-114710PC  
CURRENT APPLICATION NUMBER: PCT/US01/44737  
CURRENT FILING DATE: 2002-08-28  
PRIOR APPLICATION NUMBER: US 09/724,857  
PRIOR FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: US 60/253,672  
PRIOR FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 36  
LENGTH: 20  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: oligo (df-20)  
PCT-US01-44737-36

Query Match 0.7%; Score 20; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 6.9e+06;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 8  
PCT-US02-10873A-574

```
; Sequence 574, Application PC/TUS0210873A
; GENERAL INFORMATION:
; APPLICANT: Byrne, David J.
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Evans, Claire F.
; APPLICANT: Mah, Steven P.
; APPLICANT: Lo, David D.
; TITLE OF INVENTION: Genes Expressed in Intestinal Epithelium and Peyer's Patch M Cell
; FILE REFERENCE: 216019-109
; CURRENT APPLICATION NUMBER: PCT/US02/10873A
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/281,416
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 859
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 574
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Real-time validation FOR primer for EDDIS_87
PCT-US02-10873A-574
```

```
Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY      1845 CTCACGCGCTGTAATCCGAGC 1864
Db      1 CTCACGCGCTGTAATCCGAGC 20
```

```
RESULT 9
PCT-US02-18229-5
; Sequence 5, Application PC/TUS0218229
; GENERAL INFORMATION:
; APPLICANT: Haydock, Paul V.
; APPLICANT: U'Ren, Jack
; APPLICANT: Salsgene Corporation
; TITLE OF INVENTION: Nucleic Acid Amplification Utilizing Intermediate
; FILE REFERENCE: 018048-001711PC
; CURRENT APPLICATION NUMBER: PCT/US02/18229
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/296,812
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 10/077,383
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: (A)-12-20
; OTHER INFORMATION: homopolymer spacer sequence
; NAME/KEY: modified base
; LOCATION: (13)..(20)
; OTHER INFORMATION: a at positions 13-20 may be present or absent
PCT-US02-18229-5
```

```
Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY      1799 AAAAAAAAAAAAAAAAAAAAAA 1818
Db      1 AAAAAAAAAAAAAAAAAAAAAA 20
```

```
RESULT 10
PCT-US02-18229-6/c
; Sequence 6, Application PC/TUS0218229
; GENERAL INFORMATION:
; APPLICANT: Haydock, Paul V.
; APPLICANT: U'Ren, Jack
; APPLICANT: Salsgene Corporation
; TITLE OF INVENTION: Nucleic Acid Amplification Utilizing Intermediate
; FILE REFERENCE: 018048-001711PC
; CURRENT APPLICATION NUMBER: PCT/US02/18229
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/296,812
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 10/077,383
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: (T)-12-20
; OTHER INFORMATION: homopolymer spacer sequence
; NAME/KEY: modified base
; LOCATION: (13)..(20)
; OTHER INFORMATION: t at positions 13-20 may be present or absent
PCT-US02-18229-6
```

```
Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY      1799 AAAAAAAAAAAAAAAAAAAAAA 1818
Db      20 AAAAAAAAAAAAAAAAAAAAAA 1
```

```
RESULT 11
PCT-US02-33002-2/c
; Sequence 2, Application PC/TUS0233002
; GENERAL INFORMATION:
; APPLICANT: Serafini, Tico
; APPLICANT: Chen, Hang
; APPLICANT: Toemel, Emily
; APPLICANT: Ellis, Michael
; TITLE OF INVENTION: Methods for Nucleic Acid Amplification
; FILE REFERENCE: 2605-1-002PCT
; CURRENT APPLICATION NUMBER: PCT/US02/33002
; CURRENT FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US 10/036,860
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/305,666
; PRIOR FILING DATE: 2001-07-16
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic sequence
PCT-US02-33002-2
```

```
Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

QY      1799 AAAAAAAAAAAAAAAAAAAAAA 1818
Db      20 AAAAAAAAAAAAAAAAAAAAAA 1
```

```
RESULT 12
PCT-US02-33699-1/c
; Sequence 1, Application PC/TUS0233699
; GENERAL INFORMATION:
; APPLICANT: Prologo, LLC
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHODS TO DETECT AND/OR QUANTIFY NUCLEIC
; FILE OF INVENTION: ACID ANALYTES
; FILE REFERENCE: PRO.07/PCT
; CURRENT APPLICATION NUMBER: PCT/US02/33699
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: 60/336,432
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent version 3.2
; SEQ ID NO 1
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Nucleic Acid Probe
; NAME/KEY: misc.feature
; LOCATION: (1)..(20)
PCT-US02-33699-1

Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
|||||
Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 13
PCT-US02-35719-13
; Sequence 13, Application PC/TUS0235719
; GENERAL INFORMATION:
; APPLICANT: Wood, Linda
; APPLICANT: Wagner, Susanne
; APPLICANT: Parodi, Luis
; TITLE OF INVENTION: Single Nucleotide Polymorphisms in GH-1
; FILE REFERENCE: 00791.US1
; CURRENT APPLICATION NUMBER: PCT/US02/35719
; CURRENT FILING DATE: 2002-11-07
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent version 3.1
; SEQ ID NO 13
; LENGTH: 20
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: primer
PCT-US02-35719-13

Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1852 CTGTAATCCAGCACTTTGG 1871
|||||
Db 1 CTGTAATCCAGCACTTTGG 20

RESULT 14
PCT-US02-38281-266
; Sequence 266, Application PC/TUS0238281
; GENERAL INFORMATION:
; APPLICANT: Loren J. Miraglia
; APPLICANT: Pamela S. Nero
; APPLICANT: Brett P. Monia
; APPLICANT: Mark J. Graham
```

```
APPLICANT: Brett P. Monia
APPLICANT: Brett Koller
APPLICANT: Ming Yi Chiang
APPLICANT: Muthiah Manoharan
APPLICANT: Isis Pharmaceuticals, Inc.
; TITLE OF INVENTION: Antisense Modulation of Human mdm2 Expression
; FILE REFERENCE: ISPH-0717
; CURRENT APPLICATION NUMBER: PCT/US02/38281
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US 10/005,344
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 09/752,983
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: US 09/280,805
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: US 09/048,810
; PRIOR FILING DATE: 1998-03-26
; NUMBER OF SEQ ID NOS: 379
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 266
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
PCT-US02-38281-266

Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1858 TCCGAGCACTTTGGAGGCC 1877
|||||
Db 1 TCCGAGCACTTTGGAGGCC 20

RESULT 15
PCT-US02-38281-267
; Sequence 267, Application PC/TUS0238281
; GENERAL INFORMATION:
; APPLICANT: Loren J. Miraglia
; APPLICANT: Pamela S. Nero
; APPLICANT: Mark J. Graham
; APPLICANT: Brett P. Monia
; APPLICANT: Brett Koller
; APPLICANT: Ming Yi Chiang
; APPLICANT: Muthiah Manoharan
; APPLICANT: Isis Pharmaceuticals, Inc.
; TITLE OF INVENTION: Antisense Modulation of Human mdm2 Expression
; FILE REFERENCE: ISPH-0717
; CURRENT APPLICATION NUMBER: PCT/US02/38281
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US 10/005,344
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 09/752,983
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: US 09/280,805
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: US 09/048,810
; PRIOR FILING DATE: 1998-03-26
; NUMBER OF SEQ ID NOS: 379
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 267
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
PCT-US02-38281-267

Query Match          0.7%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred.No. 6.9e+06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



QY 1851 CCTGTAATCCGACGACTTG 1870  
Db 1 CCTGTAATCCGACGACTTG 20

Search completed: March 7, 2004, 22:42:14  
Job time : 9054 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 7, 2004, 15:07:23 ; Search time 94 Seconds  
(without alignments)  
6400.467 Million cell updates/sec

Title: US-09-966-880A-7

Perfect score: 2818  
Sequence: 1 agagaccatcatcatgta.....aaaaaaaaaaaaaaaaaa 2818

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 211252 seqs, 106750167 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27452

Minimum DB seq length: 0  
Maximum DB seq length: 20

Post-processing: Listing first 45 summaries

Database : Pending Patents\_NA.New.\*

1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	0.7	20	1 PCT-US04-04452-51	Sequence 51, Appl
C 2	20	0.7	20	1 PCT-US04-04452-1813	Sequence 1813, Ap
C 3	20	0.7	20	6 US-10-653-416-25	Sequence 25, Appl
C 4	20	0.7	20	6 US-10-213-257-23	Sequence 23, Appl
C 5	20	0.7	20	6 US-10-708-204-5116	Sequence 5116, Ap
C 6	19	0.7	20	5 US-09-574-779B-30	Sequence 30, Appl
C 7	19	0.7	20	6 US-10-708-204-5092	Sequence 5092, Ap
C 8	18	0.6	18	6 US-10-653-416-26	Sequence 26, Appl
C 9	18	0.6	18	6 US-10-708-204-3825	Sequence 3825, Ap
C 10	18	0.6	18	6 US-10-708-204-5130	Sequence 5130, Ap
C 11	18	0.6	19	1 PCT-US03-25997A-50	Sequence 50, Appl
C 12	18	0.6	19	1 PCT-US04-04452-252	Sequence 252, Appl
C 13	18	0.6	20	6 US-10-708-204-2896	Sequence 2896, Ap
C 14	18	0.6	20	6 US-10-708-204-3868	Sequence 3868, Ap
C 15	17	0.6	19	6 US-10-708-204-3854	Sequence 3854, Ap
C 16	17	0.6	19	6 US-10-708-204-5172	Sequence 5172, Ap
C 17	17	0.6	20	6 US-10-708-204-5094	Sequence 5094, Ap
C 18	16	0.6	16	6 US-10-776-099-9	Sequence 9, Appl
C 19	16	0.6	18	6 US-10-708-204-5101	Sequence 5101, Ap
C 20	16	0.6	19	1 PCT-US04-04452-253	Sequence 253, Appl
C 21	16	0.6	19	6 US-10-708-204-3840	Sequence 3840, Ap
C 22	16	0.6	19	6 US-10-708-204-5110	Sequence 5110, Ap
C 23	16	0.6	20	6 US-10-708-204-3841	Sequence 3841, Ap
C 24	16	0.6	20	6 US-10-708-204-3842	Sequence 3842, Ap
C 25	15	0.5	19	6 US-10-708-204-3842	Sequence 3842, Ap
C 26	15	0.5	19	6 US-10-708-204-4379	Sequence 4379, Ap

27	15	0.5	20	1 PCT-US04-04452-117	Sequence 117, Appl
C 28	15	0.5	20	1 PCT-US04-04452-536	Sequence 536, Appl
C 29	15	0.5	20	1 PCT-US04-04452-642	Sequence 642, Appl
C 30	15	0.5	20	6 US-10-708-204-3210	Sequence 3210, Ap
C 31	15	0.5	20	6 US-10-708-204-4471	Sequence 4471, Ap
C 32	15	0.5	20	6 US-10-708-204-5144	Sequence 5144, Appl
C 33	14	0.5	14	6 US-10-141-535-2	Sequence 2, Appl
C 34	14	0.5	19	1 PCT-US04-04452-460	Sequence 460, Appl
C 35	14	0.5	19	1 PCT-US04-04452-1367	Sequence 1367, Appl
C 36	14	0.5	20	6 US-10-769-579-2	Sequence 2, Appl
C 37	13	0.5	17	1 PCT-US03-31862-103	Sequence 103, Appl
C 38	13	0.5	17	1 PCT-US03-31862-104	Sequence 104, Appl
C 39	13	0.5	18	1 PCT-US03-32805-22	Sequence 22, Appl
C 40	13	0.5	19	1 PCT-US03-02038-164	Sequence 164, Appl
C 41	13	0.5	19	6 US-10-708-204-3140	Sequence 3140, Appl
C 42	13	0.5	20	1 PCT-US04-02344-110	Sequence 110, Appl
C 43	13	0.5	20	1 PCT-US04-04452-1667	Sequence 1667, Appl
C 44	13	0.5	20	5 US-09-574-779B-15	Sequence 15, Appl
C 45	13	0.5	20	6 US-10-484-805-9	Sequence 9, Appl

## ALIGNMENTS

RESULT 1  
PCT-US04-04452-51/C  
; Sequence 51, Application PC/TUS0404452  
; GENERAL INFORMATION:  
; APPLICANT: Bardelli, Alberto  
; APPLICANT: Parsons, Will  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Kinzler, Kenneth W.  
; APPLICANT: Vogelstein, Bert  
; TITLE OF INVENTION: TYROSINE KINASES IMPLICATED IN CANCERS  
; FILE REFERENCE: 001107.00327  
; CURRENT APPLICATION NUMBER: PCT/US04/04452  
; CURRENT FILING DATE: 2004-02-18  
; NUMBER OF SEQ ID NOS: 2191  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 51  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
PCT-US04-04452-51

Query Match 0.7%; Score 20; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1901 CAGGAGTTGAGACGAGCT 1920  
Db 20 CAGGAGTTGAGACGAGCT 1

RESULT 2  
PCT-US04-04452-1813/C  
; Sequence 1813, Application PC/TUS0404452  
; GENERAL INFORMATION:  
; APPLICANT: Bardelli, Alberto  
; APPLICANT: Parsons, Will  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Kinzler, Kenneth W.  
; APPLICANT: Vogelstein, Bert  
; TITLE OF INVENTION: TYROSINE KINASES IMPLICATED IN CANCERS  
; FILE REFERENCE: 001107.00327  
; CURRENT APPLICATION NUMBER: PCT/US04/04452  
; CURRENT FILING DATE: 2004-02-18  
; NUMBER OF SEQ ID NOS: 2191  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1813  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Homo sapiens

PCT-US04-04452-1813

Query Match 0.7%; Score 20; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1854 GTAATCCAGCACTTTGGGA 1873  
DB 20 GTAATCCAGCACTTTGGGA 1

RESULT 3

US-10-653-416-25/c  
; Sequence 25, Application US/10653416  
; GENERAL INFORMATION:  
; APPLICANT: RASHTCHIAN, AYDUB  
; APPLICANT: SCHUSTER, DAVID M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CDNA SYNTHESIS  
; FILE REFERENCE: 38266-0011  
; CURRENT APPLICATION NUMBER: US/10/653,416  
; PRIOR FILING DATE: 2003-09-03  
; PRIOR APPLICATION NUMBER: 60/407,248  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 25  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: oligonucleotide  
US-10-653-416-25

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 4

US-10-213-257-23/c  
; Sequence 23, Application US/10213257  
; GENERAL INFORMATION:  
; APPLICANT: WANG, YUAN-YUAN  
; APPLICANT: XHO, KAY-HOUI  
; APPLICANT: CHEN, SHUI-TEIN  
; APPLICANT: LIN, CHUN-CHENG  
; APPLICANT: WONG, CHI-HUEY  
; APPLICANT: LIN, CHUN-HUNG  
; APPLICANT: CHEN, HONG-SEN  
; APPLICANT: TSAI, YOW-FU  
; TITLE OF INVENTION: IMMUNO-MODULATING ANTITUMOR ACTIVITIES OF GANDERMA  
; TITLE OF INVENTION: LUCIDIM (REISHI) POLYSACCHARIDES  
; FILE REFERENCE: 4910-12  
; CURRENT APPLICATION NUMBER: US/10/213,257  
; PRIOR FILING DATE: 2002-08-06  
; PRIOR APPLICATION NUMBER: 60/310,285  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 3.2  
; SEQ ID NO 23  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: primer  
US-10-213-257-23

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
DB 20 AAAAAAAAAAAAAAAAAA 1

RESULT 5

US-10-708-204-5116/c  
; Sequence 5116, Application US/10708204  
; GENERAL INFORMATION:  
; APPLICANT: ROSETTA GENOMICS LTD  
; TITLE OF INVENTION: BIOINFORMATIALLY DETECTABLE GROUP OF NOVEL REGULATORY  
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES  
; FILE REFERENCE: 55033  
; CURRENT APPLICATION NUMBER: US/10/708,204  
; PRIOR FILING DATE: 2004-02-16  
; NUMBER OF SEQ ID NOS: 7351  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5116  
; LENGTH: 20  
; TYPE: RNA  
; ORGANISM: Homo Sapiens  
US-10-708-204-5116

Query Match 0.7%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2068 TGCCGTTGCACTCCAGCTG 2087  
DB 20 TGCCGTTGCACTCCAGCTG 1

RESULT 6

US-09-574-779B-30/c  
; Sequence 30, Application US/09574779B  
; GENERAL INFORMATION:  
; APPLICANT: VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE  
; TITLE OF INVENTION: Novel cDNAs encoding catenin-binding proteins with  
; TITLE OF INVENTION: function in signalling and/or gene regulation  
; FILE REFERENCE: 2676-4415US  
; CURRENT APPLICATION NUMBER: US/09/574,779B  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 99201543.8  
; PRIOR FILING DATE: 1999-05-17  
; NUMBER OF SEQ ID NOS: 158  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: primer FVR510F  
US-09-574-779B-30

Query Match 0.7%; Score 19; DB 5; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1846 TCAAGCCTGTAATCCAGC 1864  
DB 19 TCAAGCCTGTAATCCAGC 1

RESULT 7

US-10-708-204-5092/c  
; Sequence 5092, Application US/10708204  
; GENERAL INFORMATION:  
; APPLICANT: ROSETTA GENOMICS LTD

```
/ TITLE OF INVENTION: BIOINFORMATIALLY DETECTABLE GROUP OF NOVEL REGULATORY
/ TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: 55033
/ CURRENT APPLICATION NUMBER: US/10/708,204
/ CURRENT FILING DATE: 2004-02-16
/ NUMBER OF SEQ ID NOS: 7351
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO: 5092
/ LENGTH: 20
/ TYPE: RNA
/ ORGANISM: Homo Sapiens
US-10-708-204-5092

Query Match
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2072 GTTGACCTCAGCTGGGC 2090
DB 20 GTTGACCTCAGCTGGGC 2

RESULT 8
US-10-653-416-26/c
/ Sequence 26, Application US/10653416
/ GENERAL INFORMATION:
/ APPLICANT: RASHTCHIAN, AYCOB
/ APPLICANT: SCHUSTER, DAVID M.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CDNA SYNTHESIS
/ FILE REFERENCE: 38266-0011
/ CURRENT APPLICATION NUMBER: US/10/653,416
/ CURRENT FILING DATE: 2003-09-03
/ PRIOR APPLICATION NUMBER: 60/407,248
/ PRIOR FILING DATE: 2002-09-03
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PatentIn Ver. 3.2
/ SEQ ID NO: 26
/ LENGTH: 18
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ OTHER INFORMATION: oligonucleotide
/ FEATURE:
/ OTHER INFORMATION: this sequence may encompass 12-18 nucleotides according
/ OTHER INFORMATION: to the specification as filed
US-10-653-416-26

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1816
DB 18 AAAAAAAAAAAAAAAAAA 1

RESULT 9
US-10-708-204-3825/c
/ Sequence 3825, Application US/10708204
/ GENERAL INFORMATION:
/ APPLICANT: ROSETTA GENOMICS LTD
/ TITLE OF INVENTION: BIOINFORMATIALLY DETECTABLE GROUP OF NOVEL REGULATORY
/ TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: 55033
/ CURRENT APPLICATION NUMBER: US/10/708,204
/ CURRENT FILING DATE: 2004-02-16
/ NUMBER OF SEQ ID NOS: 7351
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO: 3825
/ LENGTH: 18
```

```
/ TYPE: RNA
/ ORGANISM: Homo Sapiens
US-10-708-204-3825

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2075 GCACTCCAGCTGGCGGA 2092
DB 18 GCACTCCAGCTGGCGGA 1

RESULT 10
US-10-708-204-5130/c
/ Sequence 5130, Application US/10708204
/ GENERAL INFORMATION:
/ APPLICANT: ROSETTA GENOMICS LTD
/ TITLE OF INVENTION: BIOINFORMATIALLY DETECTABLE GROUP OF NOVEL REGULATORY
/ TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
/ TITLE OF INVENTION: THEREOF
/ FILE REFERENCE: 55033
/ CURRENT APPLICATION NUMBER: US/10/708,204
/ CURRENT FILING DATE: 2004-02-16
/ NUMBER OF SEQ ID NOS: 7351
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO: 5130
/ LENGTH: 18
/ TYPE: RNA
/ ORGANISM: Homo Sapiens
US-10-708-204-5130

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1912 GACCAGCTGGCCACAT 1929
DB 18 GACCAGCTGGCCACAT 1

RESULT 11
PCT-US03-25997A-50/c
/ Sequence 50, Application PC/TUS0325997A
/ GENERAL INFORMATION:
/ APPLICANT: CERES, INC.
/ TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY USEFUL FOR
/ TITLE OF INVENTION: INCREASING PLANT SIZE AND INCREASING THE NUMBER AND SIZE OF LEA
/ FILE REFERENCE: 2750-1573F(PC)
/ CURRENT APPLICATION NUMBER: PCT/US03/25997A
/ CURRENT FILING DATE: 2003-08-18
/ NUMBER OF SEQ ID NOS: 50
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO: 50
/ LENGTH: 19
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Oligo primer used in the generation of labeled probes for
/ OTHER INFORMATION: hybridization from first-strand cDNA
PCT-US03-25997A-50

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 19;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1816
DB 18 AAAAAAAAAAAAAAAAAA 1

RESULT 12
PCT-US04-04452-252
```

```
; Sequence 252, Application PC/TUS0404452
; GENERAL INFORMATION:
; APPLICANT: Bardelli, Alberto
; APPLICANT: Parsons, Will
; APPLICANT: Velculescu, Victor
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: TYROSINE KINASES IMPLICATED IN CANCERS
; FILE REFERENCE: 001107.00327
; CURRENT APPLICATION NUMBER: PCT/US04/04452
; CURRENT FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 2191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US04-04452-252
```

```
Query Match          0.6%; Score 18; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2008 AGGCTGAGCAGGAGGAAAT 2025
DB      1 AGGCTGAGCAGGAGGAAAT 18
```

```
RESULT 13
US-10-708-204-2896
; Sequence 2896, Application US/10708204
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2896
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-2896
```

```
Query Match          0.6%; Score 18; DB 6; Length 20;
Best Local Similarity 72.2%; Pred. No. 3.1e+02;
Matches 13; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1851 CCTGTATCCGACGACTT 1868
DB      1 CCGUUAUCCGACGACATU 18
```

```
RESULT 14
US-10-708-204-3868/C
; Sequence 3868, Application US/10708204
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3868
; LENGTH: 20
; TYPE: RNA
; ORGANISM: Homo Sapiens
```

US-10-708-204-3868

```
Query Match          0.6%; Score 18; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2074 TGCACTCCAGCTGGGGC 2091
DB      18 TGCACTCCAGCTGGGGC 1
```

```
RESULT 15
US-10-708-204-3854/C
; Sequence 3854, Application US/10708204
; GENERAL INFORMATION:
; APPLICANT: ROSETTA GENOMICS LTD
; TITLE OF INVENTION: BIOINFORMATICALLY DETECTABLE GROUP OF NOVEL REGULATORY
; TITLE OF INVENTION: OLIGONUCLEOTIDES ASSOCIATED WITH ALZHEIMER'S DISEASE AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 55033
; CURRENT APPLICATION NUMBER: US/10/708,204
; CURRENT FILING DATE: 2004-02-16
; NUMBER OF SEQ ID NOS: 7351
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3854
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo Sapiens
US-10-708-204-3854
```

```
Query Match          0.6%; Score 17; DB 6; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2074 TGCACTCCAGCTGGGGC 2090
DB      17 TGCACTCCAGCTGGGGC 1
```

Search completed: March 7, 2004, 22:43:56  
Job time : 94 secs



```

CC sequencing consortium of the German Genome Project
CC s1 sequence also available
CC This clone is available at the RZPD in Berlin
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de
XX
FH Key Location/Qualifiers
FH
FT source
FT 1..20
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone_lib="566 (synonym: hEKd2). Vector pAMP1; host
FT X1-2b; sites NotI + SalI"
FT /dev_stage="fetal"
FT /tissue_type="kidney"
XX
SQ Sequence 20 BP; 18 A; 1 C; 0 G; 1 T; 0 other;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Db 1797 TCAAAAAAAAAAAAAAAAAA 1816
1 TCAAAAAAAAAAAAAAAAAA 20
XX
RESULT 2
ID HSM003226 standard; mRNA; EST; 20 BP.
XX
AC AL038750;
XX
SV AL038750.1
XX
DT 12-MAR-1999 (Rel. 59, Created)
DT 12-MAR-1999 (Rel. 59, Last updated, Version 1)
XX
DE Homo sapiens mRNA; EST DKFZP566M1146_r1 (from clone DKFZP566M1146)
XX
KM EST; expressed sequence tag.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
XX
RN [1]
RP 1-20
RA Otterwaelder B., Obermaier B., Mewes W., Gaassenhuber J., Wiemann S.;
RT Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
RL MIPS, Am Klopferstr 18a D-82152 Martinsried, GERMANY
XX
CC Clone from S. Wiemann, sequenced by Medigenomix within the CDNA
CC sequencing consortium of the German Genome Project
CC s1 sequence also available
CC This clone is available at the RZPD in Berlin
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de
XX
FH Key Location/Qualifiers
FH
FT source
FT 1..20
FT /db_xref="taxon:9606"
FT /mol_type="mRNA"
FT /organism="Homo sapiens"
FT /clone_lib="566 (synonym: hEKd2). Vector pAMP1; host
FT X1-2b; sites NotI + SalI"
FT /dev_stage="fetal"
FT /tissue_type="kidney"
XX

```

```

XX
SQ Sequence 20 BP; 20 A; 0 C; 0 G; 0 T; 0 other;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query Match 0.7%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Db 1799 AAAAAAAAAAAAAAAAAA 1818
1 AAAAAAAAAAAAAAAAAA 20
XX
RESULT 3
ID AM334823/c
ID LOCUS
ID DEFINITION AM334823 20 bp mRNA linear EST 31-JAN-2000
ID ACCESSION AM334823
ID VERSION AM334823.1 GI:6831180
ID KEYWORDS EST.
ID SOURCE Pneumocystis carinii
ID ORGANISM Pneumocystis carinii
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
Pneumocystidaceae; Pneumocystis.
REFERENCE
AUTHORS 1 (bases 1 to 20)
Edman, J.C., Kovacs, J. and Cushion, M.
Expressed sequence tags from Pneumocystis carinii
Unpublished (2000)
COMMENT
Contact: Staben C
School of Biological Sciences
University of Kentucky
101 Morgan Building, University of Kentucky, Lexington, KY
40506-0225, USA
Tel: 606 257 2161
Fax: 606 257 1717
Email: staben@pop.uky.edu.
FEATURES
source
location/Qualifiers
1..20
/organism="Pneumocystis carinii"
/mol_type="mRNA"
/db_xref="taxon:4754"
/lab_host="E. coli"
/clone_lib="TAGS-1"
/note="Vector: Lambda ZAP II, Site 1: EcoRI; Site 2: XhoI;
P. carinii organisms (3x10e9) from a single rat (99-1-6,
sacrificed on 3/17/99) at Cincinnati VA facilities.
Trizol extracted RNA. Oligo dT priming, standard
conditions described by vendor. Stragene. Further
details see www.uky.edu/Project/Pneumocystis/"
XX
ORIGIN
Query Match 0.7%; Score 20; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Db 1799 AAAAAAAAAAAAAAAAAA 1818
20 AAAAAAAAAAAAAAAAAA 1
XX
RESULT 4
ID CF280913
ID LOCUS
ID DEFINITION 14ETL--07-M07.D1 Rice etiolated leaf plasmid cDNA library (14ETL)
ID ACCESSION CF280913
ID VERSION CF280913.1 GI:33658299
ID KEYWORDS EST.
ID SOURCE Oryza sativa
ID ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriatoideae; Oryzae; Oryza.
XX

```

REFERENCE 1 (bases 1 to 20)  
AUTHORS Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,  
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.  
TITLE Large-scale Sequencing Analysis of Rice ESTs  
JOURNAL Unpublished (2003)  
COMMENT Contact: Nahm B.H.  
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division  
of Bioscience and Bioinformatics, Myongji University  
Yongin, Kyonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.  
Location/Qualifiers

FEATURES  
source 1..20  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:4530"  
/clone="14ETL--07-M07"  
/tissue\_type="leaf"  
/dev\_stage="14 days after germination"  
/lab\_host="E.coli DH10B"  
/clone\_lib="Rice etiolated leaf plasmid cDNA library  
(14ETL)"  
/note="Vector: pCR4-TOPO, Site 1: EcoRI; mRNA was capped  
with oligoribonucleotides and then used as templates for  
RT-PCR."

ORIGIN  
Query Match 0.7%; Score 20; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred.No. 0;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
|||||  
1 AAAAAAAAAAAAAAAAAA 20

Db 1799 AAAAAAAAAAAAAAAAAA 1818  
|||||  
1 AAAAAAAAAAAAAAAAAA 20

RESULT 5  
CF282035  
LOCUS 14ETL--09-F19.g1 Rice etiolated leaf plasmid cDNA library (14ETL)  
DEFINITION Oryza sativa CDNA clone 14ETL--09-F19, mRNA sequence.  
ACCESSION CF282035  
VERSION CF282035.1 GI:33659422  
KEYWORDS EST.  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 20)  
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,  
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.  
Large-scale Sequencing Analysis of Rice ESTs  
Unpublished (2003)  
Contact: Nahm B.H.  
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division  
of Bioscience and Bioinformatics, Myongji University  
Yongin, Kyonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.  
Location/Qualifiers

FEATURES  
source 1..20  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:4530"  
/clone="14ETL--09-F19"  
/tissue\_type="leaf"  
/dev\_stage="14 days after germination"  
/lab\_host="E.coli DH10B"  
/clone\_lib="Rice etiolated leaf plasmid cDNA library  
(14ETL)"  
/note="Vector: pCR4-TOPO, Site 1: EcoRI; mRNA was capped  
with oligoribonucleotides and then used as templates for  
RT-PCR."

/clone\_lib="Rice etiolated leaf plasmid cDNA library  
(14ETL)"  
/note="Vector: pCR4-TOPO, Site 1: EcoRI; mRNA was capped  
with oligoribonucleotides and then used as templates for  
RT-PCR."

ORIGIN  
Query Match 0.7%; Score 20; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred.No. 0;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
|||||  
1 AAAAAAAAAAAAAAAAAA 20

Db 1799 AAAAAAAAAAAAAAAAAA 1818  
|||||  
1 AAAAAAAAAAAAAAAAAA 20

RESULT 6  
CF282414  
LOCUS 14ETL--09-F19.g1 Rice etiolated leaf plasmid cDNA library (14ETL)  
DEFINITION Oryza sativa CDNA clone 14ETL--09-F19, mRNA sequence.  
ACCESSION CF282414  
VERSION CF282414.1 GI:33659801  
KEYWORDS EST.  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 20)  
Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,  
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.  
Large-scale Sequencing Analysis of Rice ESTs  
Unpublished (2003)  
Contact: Nahm B.H.  
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division  
of Bioscience and Bioinformatics, Myongji University  
Yongin, Kyonggi, Korea  
Tel: 82 31 330 6193  
Fax: 82 31 321 6355  
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.  
Location/Qualifiers

FEATURES  
source 1..20  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/cultivar="Nackdong"  
/db\_xref="taxon:4530"  
/clone="14ETL--09-F19"  
/tissue\_type="leaf"  
/dev\_stage="14 days after germination"  
/lab\_host="E.coli DH10B"  
/clone\_lib="Rice etiolated leaf plasmid cDNA library  
(14ETL)"  
/note="Vector: pCR4-TOPO, Site 1: EcoRI; mRNA was capped  
with oligoribonucleotides and then used as templates for  
RT-PCR."

ORIGIN  
Query Match 0.7%; Score 20; DB 14; Length 20;  
Best Local Similarity 100.0%; Pred.No. 0;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1799 AAAAAAAAAAAAAAAAAA 1818  
|||||  
1 AAAAAAAAAAAAAAAAAA 20

Db 1799 AAAAAAAAAAAAAAAAAA 1818  
|||||  
1 AAAAAAAAAAAAAAAAAA 20

RESULT 7  
CF299822/c  
LOCUS 7LEAF--03-P22.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza  
DEFINITION sativa CDNA clone 7LEAF--03-P22, mRNA sequence.  
ACCESSION CF299822



```

VERSION      CF299622.1  GI:33671583
KEYWORDS     EST.
SOURCE       Oryza sativa
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE    1 (bases 1 to 20)
AUTHORS      Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE        Large-scale Sequencing Analysis of Rice ESTs
JOURNAL      Unpublished (2003)
COMMENT      Contact: Nahm B.H.
              Genomics and Genetics Institute, Greengene Biotech Inc., Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea
              Tel: 82 31 330 6193
              Fax: 82 31 321 6355
              Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
              Location/Qualifiers
1..20
FEATURES
SOURCE
1..20
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="7LEAF-03-P22"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
|||||
20 AAAAAAAAAAAAAAAAAA 1

Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 8
CF301720/c 20 bp mRNA linear EST 15-AUG-2003
LOCUS 7LEAF--06-L01.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
DEFINITION sativa cDNA clone 7LEAF--06-L01, mRNA sequence.
ACCESSION  CF301720
VERSION     CF301720.1  GI:33673481
KEYWORDS   EST.
SOURCE     Oryza sativa
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE    1 (bases 1 to 20)
AUTHORS      Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE        Large-scale Sequencing Analysis of Rice ESTs
JOURNAL      Unpublished (2003)
COMMENT      Contact: Nahm B.H.
              Genomics and Genetics Institute, Greengene Biotech Inc., Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea
              Tel: 82 31 330 6193
              Fax: 82 31 321 6355
              Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
              Location/Qualifiers
1..20
FEATURES
SOURCE
1..20
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="7LEAF-07-C16"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
|||||
20 AAAAAAAAAAAAAAAAAA 1

Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 9
CF302027/c 20 bp mRNA linear EST 15-AUG-2003
LOCUS 7LEAF--07-C16.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
DEFINITION sativa cDNA clone 7LEAF--07-C16, mRNA sequence.
ACCESSION  CF302027
VERSION     CF302027.1  GI:33673788
KEYWORDS   EST.
SOURCE     Oryza sativa
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE    1 (bases 1 to 20)
AUTHORS      Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE        Large-scale Sequencing Analysis of Rice ESTs
JOURNAL      Unpublished (2003)
COMMENT      Contact: Nahm B.H.
              Genomics and Genetics Institute, Greengene Biotech Inc., Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea
              Tel: 82 31 330 6193
              Fax: 82 31 321 6355
              Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
              Location/Qualifiers
1..20
FEATURES
SOURCE
1..20
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="7LEAF-07-C16"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
|||||
20 AAAAAAAAAAAAAAAAAA 1

Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 10
CF310604

```

```

FEATURES
SOURCE
1..20
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="7LEAF-06-L01"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818
|||||
20 AAAAAAAAAAAAAAAAAA 1

Db 20 AAAAAAAAAAAAAAAAAA 1

RESULT 10
CF310604

```

LOCUS CF310604 20 bp mRNA linear EST 15-AUG-2003  
 DEFINITION ABF--05-F14.g1 ABF3-overexpressing transgenic rice plasmid cDNA  
 library (ABF) Oryza sativa cDNA clone ABF--05-F14, mRNA sequence.  
 ACCESSION CF310604  
 VERSION CF310604.1 GI:33682365  
 KEYWORDS EST.  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 20)  
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,  
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.  
 Large-scale Sequencing Analysis of Rice ESTs  
 Unpublished (2003)  
 TITLE Oryza sativa  
 JOURNAL Contact: Nahm B.H.  
 COMMENT Genomics and Genetics Institute, Greengene Biotech Inc.; Division  
 of Bioscience and Bioinformatics, Myongji University  
 Yongin, Kyeonggi, Korea  
 Tel: 82 31 321 6355  
 Fax: 82 31 321 6355  
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES  
 source  
 1..20  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /cultivar="Nackdong"  
 /db\_xref="taxon:4530"  
 /clone="ABF--05-F14"  
 /tissue\_type="leaf"  
 /dev\_stage="14 days after germination"  
 /lab\_host="E.coli DH10B"  
 /clone\_lib="ABF3-overexpressing transgenic rice plasmid  
 cDNA library (ABF)"  
 /note="Vector: PCR4-TOPO; Site 1: EcoRI; Leaf was dried  
 for 2hrs. Oligo-capped mRNA was reverse transcribed and  
 then used for PCR. mRNA was prepared from ABA-responsive  
 element binding transcription factor 3 overexpression  
 line."

ORIGIN  
 Query Match 0.7%; Score 20; DB 14; Length 20;  
 Best Local Similarity 100.0%; Pred.No. 0;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
 1 AAAAAAAAAAAAAAAAAA 20

RESULT 11  
 LOCUS CF313067  
 DEFINITION CP313067 20 bp mRNA linear EST 15-AUG-2003  
 HD--01-B02.g1 OSHDAC1-overexpressing transgenic rice plasmid cDNA  
 library (HD) Oryza sativa cDNA clone HD--01-B02, mRNA sequence.  
 ACCESSION CF313067  
 VERSION CF313067.1 GI:33684828  
 KEYWORDS EST.  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 20)  
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,  
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.  
 Large-scale Sequencing Analysis of Rice ESTs  
 Unpublished (2003)  
 TITLE Oryza sativa  
 JOURNAL Contact: Nahm B.H.  
 COMMENT Genomics and Genetics Institute, Greengene Biotech Inc.; Division  
 of Bioscience and Bioinformatics, Myongji University  
 Yongin, Kyeonggi, Korea  
 Tel: 82 31 321 6355  
 Fax: 82 31 321 6355  
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES  
 source  
 1..20  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /cultivar="Nackdong"  
 /db\_xref="taxon:4530"  
 /clone="HD--01-B02"  
 /tissue\_type="callus"  
 /dev\_stage="proliferated callus on 2N6 media for 2 weeks"  
 /lab\_host="E.coli DH10B"  
 /clone\_lib="OSHDAC1-overexpressing transgenic rice plasmid  
 cDNA library (HD)"  
 /note="Vector: PCR4-TOPO; Site 1: EcoRI; Callus was  
 treated with ABA(20um) for 1hr. Oligo-capped mRNA was  
 reverse transcribed and then used for PCR. mRNA was  
 derived from rice Histone Deacetylase overexpression  
 line."

ORIGIN  
 Query Match 0.7%; Score 20; DB 14; Length 20;  
 Best Local Similarity 100.0%; Pred.No. 0;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AAAAAAAAAAAAAAAAAA 1818  
 1 AAAAAAAAAAAAAAAAAA 20

RESULT 12  
 LOCUS CF313569  
 DEFINITION CP313569 20 bp mRNA linear EST 15-AUG-2003  
 HD--01-L22.b1 OSHDAC1-overexpressing transgenic rice plasmid cDNA  
 library (HD) Oryza sativa cDNA clone HD--01-L22, mRNA sequence.  
 ACCESSION CF313569  
 VERSION CF313569.1 GI:33685330  
 KEYWORDS EST.  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1 (bases 1 to 20)  
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,  
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.  
 Large-scale Sequencing Analysis of Rice ESTs  
 Unpublished (2003)  
 TITLE Oryza sativa  
 JOURNAL Contact: Nahm B.H.  
 COMMENT Genomics and Genetics Institute, Greengene Biotech Inc.; Division  
 of Bioscience and Bioinformatics, Myongji University  
 Yongin, Kyeonggi, Korea  
 Tel: 82 31 321 6355  
 Fax: 82 31 321 6355  
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES  
 source  
 1..20  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /cultivar="Nackdong"  
 /db\_xref="taxon:4530"  
 /clone="HD--01-L22"  
 /tissue\_type="callus"  
 /dev\_stage="proliferated callus on 2N6 media for 2 weeks"  
 /lab\_host="E.coli DH10B"  
 /clone\_lib="OSHDAC1-overexpressing transgenic rice plasmid  
 cDNA library (HD)"  
 /note="Vector: PCR4-TOPO; Site 1: EcoRI; Callus was  
 treated with ABA(20um) for 1hr. Oligo-capped mRNA was  
 reverse transcribed and then used for PCR. mRNA was  
 derived from rice Histone Deacetylase overexpression  
 line."

```

ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1799 AAAAAAAAAAAAAAAAAA 1818
Db      20 AAAAAAAAAAAAAAAAAA 1

RESULT 13
LOCUS      CF319133/c      20 bp      mRNA      linear      EST 15-AUG-2003
DEFINITION HD--09-113.b1 OSHDAC1-overexpressing transgenic rice plasmid cDNA
ACCESSION  CF319133
VERSION     CF319133.1 GI:33690894
KEYWORDS
ORGANISM   Oryza sativa
SOURCE     Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharctoidae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
          Genomics and Genetics Institute, Greengene Biotech Inc.; Division
          of Bioscience and Bioinformatics, Myongji University
          Yongin, Kyeonggi, Korea
          Tel: 82 31 330 6193
          Fax: 82 31 321 6355
          Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
          Location/Qualifiers
            1..20
            /organism="Oryza sativa"
            /mol_type="mRNA"
            /cultiVar="Nackdong"
            /db_xref="taxon:4530"
            /clone="HD-09-113"
            /tissue_type="callus"
            /dev_stage="proliferated callus on 2N6 media for 2 weeks"
            /lab_host="E.coli DH10B"
            /clone_lib="OSHDAC1-overexpressing transgenic rice plasmid
            cDNA library (HD)"
            /note="Vector: PCR4-TOP0; Site_1: EcoRI; Callus was
            treated with ABA(20um) for 1hr. Oligo-capped mRNA was
            reverse transcribed and then used for PCR. mRNA was
            derived from rice Histone Deacetylase overexpression
            line."

FEATURES
source

```

```

ORGANISM   Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharctoidae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
          Genomics and Genetics Institute, Greengene Biotech Inc.; Division
          of Bioscience and Bioinformatics, Myongji University
          Yongin, Kyeonggi, Korea
          Tel: 82 31 330 6193
          Fax: 82 31 321 6355
          Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
          Location/Qualifiers
            1..20
            /organism="Oryza sativa"
            /mol_type="mRNA"
            /cultiVar="Nackdong"
            /db_xref="taxon:4530"
            /clone="HD-13-B05"
            /tissue_type="callus"
            /dev_stage="proliferated callus on 2N6 media for 2 weeks"
            /lab_host="E.coli DH10B"
            /clone_lib="OSHDAC1-overexpressing transgenic rice plasmid
            cDNA library (HD)"
            /note="Vector: PCR4-TOP0; Site_1: EcoRI; Callus was
            treated with ABA(20um) for 1hr. Oligo-capped mRNA was
            reverse transcribed and then used for PCR. mRNA was
            derived from rice Histone Deacetylase overexpression
            line."

ORIGIN
Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1799 AAAAAAAAAAAAAAAAAA 1818
Db      20 AAAAAAAAAAAAAAAAAA 1

RESULT 15
LOCUS      CF328565      20 bp      mRNA      linear      EST 18-AUG-2003
DEFINITION NACL--03-114.b1 Rice callus plasmid cDNA library (NACL) Oryza
ACCESSION  CF328565
VERSION     CF328565.1 GI:33805376
KEYWORDS
ORGANISM   Oryza sativa
SOURCE     Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharctoidae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
          Genomics and Genetics Institute, Greengene Biotech Inc.; Division
          of Bioscience and Bioinformatics, Myongji University
          Yongin, Kyeonggi, Korea
          Tel: 82 31 330 6193
          Fax: 82 31 321 6355
          Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
          Location/Qualifiers
            1..20
            /organism="Oryza sativa"
            /mol_type="mRNA"

```

```

/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="NACL-03-114"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 30 days"
/lab_host="E.coli DH10B"
/clone_lib="Rice callus plasmid cDNA library (NACL)"
/notes="Vector: PCR4-TOPO; Site_1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

```

ORIGIN

```

Query Match      0.7%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1799 AAAAAAAAAAAAAAAAAA 1818
    |||||
Db 1 AAAAAAAAAAAAAAAAAA 20

```

Search completed: March 7, 2004, 20:07:28  
 Job time : 7092 secs